*eJournal of Tax Research* (2023) vol. 20, no. 2, pp. 102-167

# The geography of corporate tax avoidance

Yuan Helen Ping,[*] Hai Wu[**] and Xiu-Ye Zhang[***]

***Abstract***

We empirically examine the relation between firms' headquarters location and their level of tax avoidance. Employing multiple measures of tax avoidance, we consistently find significant location fixed effects on firms' tax behaviour across different geographic areas in the US, after controlling for firm fixed effects, time-varying firm characteristics and state income tax rates. Additional analyses show that location fixed effects are more pronounced for firms that have been located in an area for a longer period and that have lower geographic diversification. We then explore a range of regional characteristics as determinants of location fixed effects and find some evidence that location-specific resources and risks factors, but not cultural factors, are associated with time-invariant differences in corporate tax avoidance across regions. Our study has important practical implications for tax authorities, suggesting that tax enforcement, education, and inspections should be tailored to take account of firms' geographical location.

**Keywords**: tax avoidance; geographic area; location fixed effects; location-based characteristics

[*] Research School of Accounting, College of Business and Economics, Australian National University, Canberra, ACT, 2600, Australia. Email: yuan.ping@anu.edu.au (corresponding author).
[**] Research School of Accounting, College of Business and Economics, Australian National University, Canberra, ACT, 2600, Australia. Email: steven.wu@anu.edu.au.
[***] Research School of Accounting, College of Business and Economics, Australian National University, Canberra, ACT, 2600, Australia. Email: xiu-ye.zhang@anu.edu.au.

## 1. INTRODUCTION

Geographical location affects individual decision-making, leading to uneven distributions of economic and social outcomes such as innovations, health, crime and violence, as well as pro- and anti-social behaviour (e.g., Shaw & McKay, 1942; Land, McCall & Cohen, 1990; Jaffe, Trajtenberg & Henderson, 1993; Glaeser, Sacerdote & Scheinkman, 1996; Weisburd, 2015). While social scientists use location as a common unit of analysis to study individual behaviour, business studies have just begun to explore how spatial variations affect corporate behaviour. Parsons, Sulaeman and Titman (2018) take an initial step towards documenting significant differences in firms' financial misconduct rates across cities in the US. In this study, we examine whether there is a spatial variation in corporate tax avoidance policy.

Tax avoidance can be broadly defined as strategies that reduce a firm's tax burden (Dyreng, Hanlon & Maydew, 2008; Hanlon & Heitzman, 2010), not necessarily indicating any corporate wrongdoing. However, the boundary between aggressive tax schemes and breaches of tax obligations can be crossed without a clear distinction, and the extreme forms of tax avoidance, including tax non-compliance, evasion, and sheltering, are of great interest to tax authorities (Hanlon & Heitzman, 2010). Prior studies have documented several space-related determinants of corporate tax avoidance, including information flow within corporate group (Su, Li & Ma, 2019; Chen et al., 2022) and between firms and regulators (Kubick et al., 2017), as well as social characteristics surrounding firms' headquarters (Hasan et al., 2017a). Our study makes a novel contribution by directly investigating if firms' location affects their level of tax avoidance, and thus provides evidence on whether geography is a potential factor explaining the observed persistent variation in corporate tax avoidance (Dyreng et al., 2008). We then investigate the extent to which the estimated time-invariant location fixed effects can be explained by a range of observable regional characteristics, including spatial differences in local information/human resources, risk and cultural factors.

Following prior literature (John, Knyazeva & Knyazeva, 2011; Arena & Dewally, 2012; Kedia & Rajgopal, 2011; Parsons et al., 2018), we argue that the spatial patterns that systematically affect corporate tax avoidance behaviour can arise through three channels. First, firms located in the same region are likely to obtain and share similar information and resources. Variations in access to information and resources (e.g., tax expertise) across locations can affect firms' capability to adopt complex tax strategies, resulting in different levels of tax avoidance across space. Second, firms that are located closely to one another are exposed to similar location-induced risk attitudes. These location-induced risk attitudes may alter the incentives and behaviour of managers when making corporate tax decisions. Third, firms located in the same community could be affected by similar social and ethical norms. Managers are likely to behave in a way consistent with the beliefs and behaviours of their social peers. Taking these arguments together, we predict that geographical differences in those factors could lead to spatial variation in corporate tax decisions.

We estimate location fixed effects on four measures of corporate tax avoidance to encompass an entire continuum of tax planning strategies that reduce tax payments. Following Dyreng, Hanlon and Maydew (2010), we first use two standard measures to capture tax avoidance broadly: firms' effective tax rate (ETR) and cash effective tax rate (CETR). Given that location-based factors can influence both firms' use of questionable tax strategies and legitimate means to reduce tax burdens, it is important to study broad-based tax avoidance measures. We then look at the more aggressive end of the tax avoidance continuum by following prior studies which suggest that extreme unethical behaviours demonstrate high levels of geographic clustering (Glaeser et al., 1996; Parsons et al., 2018). Our third proxy for tax avoidance is the publicly disclosed tax reserves, i.e., the unrecognised tax benefits (hereafter UTB), made available through Financial Interpretation No. 48 (FIN 48). Lisowsky, Robinson and Schmidt (2013) suggest that the UTB is the most robust proxy for tax shelters. Last, we employ the long-run CETR measure to capture firms' ability to maintain aggressive tax positions for a much longer term of five years (CETR5, Dyreng et al., 2008) as we conjecture that the long-run CETR reflects persistent tax avoidance properties that are closely related to our objective of identifying time-invariant location fixed effects.

Our study focuses primarily on the US Metropolitan Statistical Area (MSA) in which firms' headquarters are located. We focus on the location of headquarters since the overall tax strategy of a firm is typically formulated and executed by the top management team (Hasan et al., 2017a; Dyreng et al., 2010). We regress the sample firms' ETR, CETR, UTB, and CETR5 respectively on location fixed effects after controlling for year and firm fixed effects as well as a set of time-varying firm-level characteristics and state corporate income tax rates. We obtain the location fixed effects by estimating the MSA-specific coefficients. There are two important features of our empirical design. First, we specifically control for firm fixed effects, thereby testing whether the within-firm variation in tax avoidance is systematically associated with the locations of firms' headquarters. Thus, our primary identification relies on observed changes of firm headquarters locations. Second, we specifically control for state-level corporate tax rates. Therefore, we only estimate tax-code unrelated variation in location effects on corporate tax behaviour.

Our empirical results show that the location of firm headquarters has a statistically and economically significant effect on the level of tax avoidance of firms. The estimated MSA fixed effects on ETR, CETR, UTB, and CETR5 are all jointly significant, regardless of whether tested individually or in the presence of other fixed effects. We find that the explanatory power of models including MSA fixed effects is in between that of those including year fixed effects and those including industry fixed effects. Moreover, the distribution of the MSA fixed effect coefficients reveals large and significant differences in tax avoidance behaviour across locations. In particular, moving between the top and bottom quartiles of MSAs results in an approximately 21% (14%) swing in CETRs (ETRs). Further evidence on the positively correlated relationship between the estimates of location fixed effects in neighbouring MSAs, and the over-time stability of location fixed effects affirms the validity of our identification on the location fixed effects. Furthermore, we conduct a number of robustness tests, such as, including loss firms in the sample, using alternative geographic units including state, county and zip code, splitting the sample to only firms that have changed headquarters location or those that have not changed location, and including executive fixed effects and controls of corporate governance and executive compensation, all of which consistently show significant effects of firm location on all of our tax avoidance variables, i.e., ETR, CETR, UTB, and CETR5. Our cross-sectional analyses by partitioning the sample based on length of time in the location and geographic diversification corroborate the main findings.

After establishing significant location fixed effects, we investigate whether these effects are associated with observable regional characteristics. Specifically, we regress the vectors of estimated MSA fixed effect coefficients obtained from the ETR, CETR, UTB, and CETR5 models on that particular MSA's: (1) information and resource factors (proxied by workforce population, education level, external accounting and finance expertise from audit firms, and geodesic proximity to Internal Revenue Service (IRS) local office); (2) economic, regulatory, and behavioural risk attitudes (as captured by average personal wage, GDP per capita, proximity to IRS, and weather pattern); and (3) social and cultural environment (including crime rates and religiosity).[1] We find evidence suggesting that locations with higher average education level and longer average daily sunlight are associated with higher corporate cash effective tax rates and lower reported tax reserves. Thus, the geographic variation in corporate tax avoidance is associated with information/resource and risk factors, while cultural factors exhibit little explanatory power. Furthermore, the low explanatory power of these regressions suggests that most of the common factors that would be able to explain location-specific tax avoidance are yet to be identified.

This study makes several contributions. First, it underscores the importance of firm location to corporate tax decisions. Prior research suggests that tax avoidance behaviour is related to several location-based characteristics (Kubick et al., 2017; Hasan et al., 2017a). Su et al. (2019) and Chen et al. (2022) find that tax avoidance is related to corporate geographical dispersion and intra-group geographic proximity. We extend this stream of literature by investigating and quantifying the overall location fixed effects on tax avoidance. This is a critical step towards a better understanding of the spatial differences in the tax avoidance undertaken by firms. Our results also suggest that little is known about the key determinants of spatial variation in corporate tax avoidance. Moreover, we do not find robust evidence confirming the effect of several previously examined location-based factors (e.g., social and culture factors) in our setting, suggesting that these results may be sensitive to research design and subject to possible time-variant omitted variables. These observations highlight the need for further research examining the effects of location-based factors on corporate tax avoidance.

Second, this research adds to the literature investigating the effects of corporate location on important corporate decisions. Prior studies document that geographic factors have an influence on individual behaviours (e.g., Land et al., 1990; Sampson, Raudenbush & Earls, 1997; Glaeser et al., 1996; Baller et al., 2001). Recent studies have emerged investigating their influence on corporate opportunistic reporting, which is related to local cultural characteristics (Parsons et al., 2018). We add to this research effort by examining whether there is a spatial variation in corporate tax avoidance activities. Financial misconduct is a clearly illegal decision, while our study shows that corporate location also matters for tax avoidance, which is a more common and recurring decision for firms. In the tax avoidance context, we find that local social and cultural characteristics are not the main determinants of spatial variation. Overall, our evidence not only indicates a more general inference regarding location effect on corporate decisions, but also suggests that different corporate policies may be driven by different location-based factors.

[1] Several geographic factors are related to more than one channel. We explain these factors and the different predictions under the different channels in section 2.3.

Third, this study has important practical implications for tax authorities. Empirical evidence from the literature that links spatial distributions to location-based economic, social, and enforcement characteristics (Land et al., 1990; Sampson et al., 1997; Glaeser et al., 1996; Baller et al., 2001) has guided resource-constrained law enforcement agencies to develop spatially-based enforcement programs (Sherman & Weisburd, 1995), as well as programs focusing on social controls and community relationships (Skolnick & Bayley, 1986). Similarly, resource-constrained tax authorities need to identify high-risk targets for tax auditing and can adopt similar tactics by using spatially-based programs to target aggressive tax behaviour. The findings of our study suggest that tax authorities could place greater reliance on location-based analysis to identify aggressive corporate taxpayers headquartered in high-risk locations. Tax enforcement, education, and inspections should happen disproportionally across space.

In the next section, we provide a review of relevant studies and develop hypotheses to address the research questions. Section 3 describes the sample selection process and presents descriptive statistics. We discuss the main results from location fixed effects in section 4 together with the robustness tests and the cross-sectional tests. Section 5 reports the relationship between location fixed effects and the observable geographical characteristics. We conclude in Section 6.

## 2. LITERATURE REVIEW AND HYPOTHESIS DEVELOPMENT

### 2.1 Prior literature on tax avoidance

Hanlon and Heitzman (2010) describe the term ‘tax avoidance’ as a continuum of tax strategies, with tax reductions that are squarely in compliance with the tax code at one end, and strategies such as so-called tax aggressiveness and tax sheltering residing closer to the other end of the continuum. We expect location-based factors can affect the entire spectrum of tax strategies. Thus, this study examines the broadly defined corporate tax avoidance, which encompasses anything that reduces a firm’s tax burden, either in the short or long term (Dyreng et al., 2008), as well as the narrowly defined concept of tax sheltering, a more aggressive type of tax avoidance.

Prior studies show that the level of tax avoidance is significantly related to a wide range of firm-level factors, including size (Zimmerman, 1983; Gupta & Newberry, 1997), profitability (Gupta & Newberry, 1997; Richardson & Lanis, 2007), life cycle (Hasan et al., 2017b), ownership structure (Chen et al., 2010), asset mix (Gupta & Newberry, 1997; Richardson & Lanis, 2007), and foreign operations (Rego, 2003). Tax avoidance is also linked to management styles (Dyreng et al., 2010), governance structures (Desai & Dharmapala, 2006; Abdul Wahab and Holland, 2012), incentive compensation (Armstrong, Blouin & Larcker, 2012), and executives’ personal tax behaviour (Hjelström et al., 2020).

Another line of studies shows that firms’ tax avoidance decisions are also shaped by external factors such as tax enforcement (Hoopes, Mescall & Pittman, 2012), government policies (Clausing, 2009), and tax professionals (McGuire, Omer & Wang, 2012b). Because many external environmental factors vary systematically across space, several papers (Kubick et al., 2017; Hasan et al., 2017a) find that factors associated with corporate location are possible determinants of firms’ tax avoidance decisions. Kubick et al. (2017) show that firms’ proximity to the regulator appears to have a significant influence on tax avoidance. Hasan et al. (2017a) provide evidence that the social environment of a firm matters for its tax avoidance decisions. This study extends prior literature to examine whether firms' geographic locations help explain persistent variation in tax avoidance that prior research has found across firms (Dyreng et al., 2008).

Our study is related to two recent studies (Chen et al., 2022, and Su et al., 2019) which examine the relation between geographical location and tax avoidance. Su et al. (2019) finds a negative effect of geographic dispersion on tax avoidance as the result of increased difficulty in intra-firm internal control and corporate governance. Chen et al. (2022), on the other hand, show a positive effect of geographic proximity between parent companies and subsidiaries on tax avoidance through intra-group income shifting to low-tax jurisdictions at lower costs. Both studies focus on the proximity/dispersion of firms or units in the same corporate group and how geography facilitates the internal information flow and coordination. Our study differs from their work as we focus our attention on investigating whether geography is an important factor influencing the tax avoidance of proximate firms regardless of whether they are in the same corporate structure. We also examine the relation of corporate tax avoidance with location-based resource, economic, risk and culture factors, rather than the geography-related internal information and coordination effects on corporate tax strategies.

## 2.2 The effects of firm location on corporate tax avoidance

The geographical effects on corporate decision-making have been well documented in many settings. For example, investors have stronger preferences for geographically local investment (e.g., Baik, Kang & Kim, 2010; Doukas & Pantzalis, 2003); auditors provide higher-quality audit services to local clients (e.g., Choi et al., 2012); analysts are more accurate in forecasting the performance of geographically proximate firms (e.g., Malloy, 2005); and regulators are more likely to investigate firms that are located closer to their local offices (e.g., Kedia & Rajgopal, 2011). The systematic geographic differences are also present in firm performance and corporate decisions, including innovative activities (Audretsch & Feldman, 1996; Jaffe et al., 1993), dividend decisions (John et al., 2011), investor clientele (Arena & Dewally, 2012), and corporate misconduct (Parsons et al., 2018; Kedia & Rajgopal, 2011). Based on this stream of literature, we believe there are at least three possible reasons why we may observe geographic variations in corporate tax decisions.

First, firm locations may be correlated with advantages in information and resources. We label this explanation as the 'information and resource channel'. Multiple studies have shown that firms located in different regions have different information sets about a range of stakeholders, including investors, creditors, and regulators (John et al., 2011; Arena & Dewally, 2012; Kedia & Rajgopal, 2011; Kubick et al., 2017). Acquisition of sensitive and informal ('soft') information is more likely to be facilitated by repeated social, civic, and business interactions and close spatial proximity between corporate managers and stakeholders (Audretsch & Stephan, 1996; Baik et al., 2010; Choi et al. 2012; Doukas & Pantzalis, 2003; Malloy, 2005). Firms located in the same area are likely to share similar information, hence making similar decisions. Furthermore, firms located in different areas may have access to different levels of resources. These resources include high-quality workforces, high-quality tax advisors, and better local infrastructure. This regional variation in business information and resources may account for the observed spatial differences in corporate behaviours and strategies.

Second, firms that are closely located to one another are exposed to similar regional risks, which alter the incentives and behaviour of managers when making risky tax avoidance decisions. We label this explanation as the 'risk and attitude channel'. For example, firms may face more regulatory risks if they are geographically proximate to regulatory and enforcement bodies, such as the IRS, because those resource-constrained regulators are more likely to investigate firms located closer to their offices (Kedia & Rajgopal, 2011; Kubick et al., 2017). These spatial differences in enforcement and/or detection efforts may make firms perceive the probability of regulatory examination and the associated cost of committing a violation differently. Differences among firms in their cost-benefit analyses of a violation would in turn affect whether they decided to break the rules or respect them. In addition to enforcement, it is possible that firms' decisions to engage in certain corporate activities may be influenced by other regional factors that alter people's attitude towards risk, including those related to overall economic wellbeing and weather-induced psychological conditions in a particular region. The effects that risk attitudes have on corporate behaviours and strategies may partially explain why firms that are located close to each other behave similarly and those that are distant from each other behave differently.

Third, firm locations may be correlated with the cultural and ethical norms of local communities, which may differ widely across space (Parsons et al., 2018; Hasan et al., 2017a). This explanation is labelled as the 'social channel'. Different social environments result in different shared common beliefs and attitudes concerning the civic duty and acceptable behaviours of the residents, including corporations, located in the community. Inevitably, managers of firms could be affected by the local community's culture, and they are likely to behave in a way that is consistent with the beliefs and behaviours of their social peers. The anticipated reputational costs and social sanctions associated with norm-deviant behaviours inhibit managers from going against the expectations of the community when running corporations. Since corporate tax avoidance is a contentious issue that affects a firm and its top-level managers' reputations in their surrounding community, managers are likely to take account of local social norms in making tax avoidance decisions (Gallemore, Maydew & Thornock, 2014).

Taking the above arguments together, we posit that the level of tax avoidance activities undertaken by a given firm will depend on where the firm is located. This study attempts to estimate the fixed effects that firms' locations have on corporate tax avoidance practices.

*Hypothesis*: Firms' locations have significant fixed effects on corporate tax avoidance.

### 2.3 Location-based characteristics and location fixed effects

We further investigate whether location fixed effects are systematically associated with various location-based characteristics. In particular, we focus on measuring the three potential channels outlined above. We derive several factors that aim to capture these characteristics. As the current analysis is exploratory in nature, we provide arguments without formally developing hypotheses for each factor considered. We note that several factors are linked to more than one explanation outlined above. In these cases, we explore how different channels lead to different predictions of the relationship between these factors and tax avoidance.

We begin with geographic proximity to the IRS territory manager's office. This factor relates to both the information channel and the risk channel with different predictions. On the one hand, the proximity to the IRS can provide information advantages to corporate taxpayers regarding local IRS enforcement methods and priorities (Kubick et al., 2017). This superior information could enable nearby companies to engage in more tax avoidance activities, as they could tailor their tax planning strategies to take advantage of additional information about regulatory scrutiny. On the other hand, because tax authorities have limited resources, enforcement risks decrease with geographical distance to the tax authorities. Firms anticipating a higher likelihood of IRS examination are less likely to take aggressive tax positions. Kubick et al. (2017) show that the effect of proximity to the IRS on corporate tax avoidance is more consistent with the information advantage argument.

Second, we consider local workforce education level. We argue that this factor also reflects both the information channel and the risk channel. Call et al. (2017) find that high-quality local workforces are associated with reduced errors in reporting. Education is particularly important for specialised work such as tax management, where employees are required to be familiar with complex tax codes and tax avoidance strategies. Thus, we expect highly educated employees to be better at coping with such complexity, which could help firms reduce their tax payments. Education may also affect workers' attitude towards tax avoidance risk. Highly educated workers may better understand the risk associated with tax avoidance, leading them to better comply with the tax code. Moreover, there is evidence that education is positively associated with risk aversion (Jung, 2015). Thus, the risk channel predicts a negative association between regional education level and tax avoidance. Therefore, whether education can be positively or negatively related to tax avoidance is an empirical question.

Third, we explore local economic conditions, proxied by average personal wages and gross domestic product (GDP) per capita. The risk-based explanation suggests that local economic conditions can affect firms' tax avoidance in two ways. On the one hand, firms located in less developed areas may be exposed to more financial risks, and this could provide managers with more incentives to avoid tax. On the other hand, less developed areas have a much larger need than well-developed areas for more tax revenue to invest in basic public goods and services, and this could result in local tax authorities' tightened monitoring and detection of any tax avoidance behaviours of corporate taxpayers. We note that local economic factors also indicate the resource level of a specific area.

The fourth factor we examine is local weather patterns. The impact of sunlight and cloud on economic activity and corporate decisions has been documented by several previous studies (Kamstra, Kramer & Levi, 2003; Goetzmann & Zhu, 2005; Goetzmann et al., 2015). Chen et al. (2019) argued that variations in weather conditions induce mood fluctuations, and mood affects individual and corporate decision-making. Specifically, reduced daily sunlight and increased cloud cover trigger managers' negative mood and pessimism which are more likely to lead them to perceive greater cash flow risks. Therefore, they may engage in more aggressive tax avoidance to preserve internally generated cash flows. We predict a negative association between tax avoidance and local daily sunlight.

Fifth, we investigate whether local tax expertise may affect location fixed effects of tax avoidance. Our variable is the presence of 'Big 4' audit firms' local offices. Audit firms provide tax consulting services to help clients with tax planning. Auditors' tax-specific expertise is associated with greater tax avoidance (McGuire et al., 2012b). Firms that are located in areas with local Big 4 audit offices have easier access to high quality tax consulting services. This local resource can assist firms with the design and implementation of tax strategies. Thus, the information and resource channel suggests a positive relationship between the presence of Big 4 audit firms' local offices and tax avoidance.

Sixth, we investigate local criminal culture. The average crime rates of a region indicate local residents' general attitudes toward extreme and illegal behaviours (Parsons et al., 2018), which could be related to the more aggressive tax sheltering activities of some firms. Given that a substantial number of tax avoidance activities are perfectly legal and therefore not subject to formal enforcement, managers' decisions to engage in those activities could be affected by whether the local community at large views tax avoidance as misbehaviour and thereby infers any wrongdoing on the part of the firms. If a community has a widely shared belief that it is inappropriate and unacceptable for all residents, including corporate residents, to avoid paying any tax, it is likely that firms located in that community and their managers will bear higher social and reputational costs from engaging in any type of aggressive tax strategy, and this may alter their incentives regarding tax planning and avoidance (Hasan et al., 2017a).

Last, McGuire, Omer and Sharp (2012a) find that firms headquartered in MSAs with religiously adherent residents have fewer incidents of accounting fraud, and they conclude that religion acts as a substitute for regulatory monitoring. Thus, religiosity is an important component of local ethical culture. It is likely that firms located in these religious regions will be less tax aggressive, as a means of reducing the expected associated social costs.

## 3. SAMPLE SELECTION AND DESCRIPTIVE STATISTICS

Our primary sample begins with all publicly listed firms with headquarters located in the US for the years 1994 to 2017. We obtain financial information from COMPUSTAT. Financial firms (SIC 6000-6999) and firms in the highly regulated industries (SIC 4400-5000) are excluded due to inherent regulatory and institutional differences. Due to data availability constraints, we construct different samples for the different measures of tax avoidance. There are four measures of tax avoidance, i.e., ETR, CETR, UTB, and CETR5. We elaborate upon the specific measurement in the next section and Table 7 (Variable Definitions, Appendix). For the analysis of ETR and CETR, we use the full sample period of 1994-2017. As firm-level UTB data only became publicly available for fiscal years beginning after December 2006, the UTB sample period is from 2007 to 2017. To ensure a valid economic interpretation of the ETR measures, we follow prior literature to exclude firm-year observations with negative pre-tax book income, negative income tax expense, and negative cash tax paid (Dyreng et al., 2010; Kubick et al., 2017). Requiring the ETRs to be within the range of 0 to 1 ensures valid interpretation of our results (Dyreng et al., 2010). In the final sample, we retain only firm-year observations that have the requisite data to construct all variables in our main analysis. We obtain historical headquarters addresses (zip codes) from firms' 10-K filings in the US Securities and Exchange Commission (SEC) EDGAR database. Firm-year observations with missing or invalid location data are also excluded. To eliminate the noise from the effects of MSAs in which only small numbers of firms are headquartered, we require each MSA to have at least two sample firms in each fiscal year.

The resulting primary sample for ETR/CETR comprises 29,293 firm-year observations, corresponding to 5,197 distinct firms with headquarters located in 218 MSAs, 50 states, and 557 counties, and having 2,769 zip codes. The UTB sample has 18,925 observations in 158 MSAs, and the CETR5 sample has 20,707 observations in 164 MSAs. We report this sample selection process in Panel A, Table 1. Since we also aim to examine the possible geographic factors that may be associated with location effects, the number of MSAs in the sample is further reduced to 120 for the ETR/CETR sample (86 for the UTB sample and 77 for the CETR5 sample) due to missing data for some geography-based explanatory variables. Panel A in Table 1 (Sample Selection, Appendix) provides a description of the sample for this part of the examination.

Table 2 (Descriptive Statistics, Appendix) presents descriptive statistics for the variables in our samples. Following Kubick et al. (2017), all continuous tax and control variables are winsorised at the 1% and 99% levels to mitigate the influence of outliers. The ETR variable is measured as the ratio of total tax expense to pre-tax income. The mean ETR in the sample is 31.5%, with a median of 33.5%. The CETR variable is measured as the ratio of cash taxes paid to pre-tax income. The mean CETR is 25.9% and the median is 24.1%. The UTB variable is measured as the ratio of the end-of-year unrecognised tax benefits to total assets. The mean UTB is 1.2% and the median is 0.4%. The CETR5 variable is the long-run CETR, computed as the sum of cash tax paid over a five-year period divided by pre-tax income over the same period.[2] The mean CETR5 is 25.6% and the median is 26.8%. These statistics are comparable to those in prior studies (Dyreng et al., 2010; Kubick et al., 2017; Hasan et al., 2017a).

Both the mean and median values of ETR and CETR are lower than the US federal corporate tax rate (35%) during the sample period due to the numerous deduction provisions in the tax code. The observation that CETR is lower than ETR is consistent with prior research (Mills, 1998; Dyreng et al., 2010), which suggests that firms generally have lower taxable income than pre-tax book income. The distributions of CETR and CETR5 are comparable, indicating that a lot of tax planning strategies cannot be implemented within a short time frame and thus require multi-year engagement from the firm (Hoopes et al., 2012).

The distribution of these tax measures reveals that there is significant variation in tax avoidance. At the 25th percentile, CETR is only 10.2%, but is 34.9% at the 75th percentile. This means that there are many firms in the sample that have successfully engaged in substantial tax avoidance, but at the same time, there are also a large proportion of firms that appear to be engaging in little or no tax avoidance, and some even pay taxes in excess of the statutory tax rate of 35%. The evidence suggests that not all firms take advantage of tax avoidance opportunities to reduce their tax payments.

We also present a number of other variables in Table 2 that are used to capture time-varying characteristics of the sample firms, and differences in state corporate income tax rates. All of them are defined in Panel A of Table 7 (Variable Definitions, Appendix). The summary statistics for those control variables are in the range of those in the extant literature (e.g., Dyreng et al., 2010; Kubick et al., 2017; Hasan et al., 2017a).

[2] To ensure a meaningful interpretation of the long-run effective tax rate measure, there needs to be a positive denominator. Following Dyreng et al. (2008), we require our sample firms' pre-tax income to be positive when summed over the five-year period (t-4 to t).

# 4. MAIN RESULTS: LOCATION FIXED EFFECTS

## 4.1 Primary model specification

We first examine two widely used standard measures: ETR and CETR. It is important to note that neither of these measures is able to capture tax avoidance perfectly. ETR captures only permanent tax strategies and is affected by managerial decisions related to financial reporting for income taxes (De Simone et al., 2020). In contrast, CETR captures both permanent and temporary tax strategies, but is more volatile as it can be affected by any action that reduces a firm's explicit tax liability (Dyreng et al., 2010). Despite those limitations, ETR and CETR have been employed broadly to detect tax avoidance because they can be calculated easily using public financial statement information. The third measure, UTB, captures the tax practices designed principally to avoid or evade taxes (Lisowsky et al., 2013). UTB reflects more aggressive tax avoidance strategies which are arguably riskier and less ethical. Lastly, we examine location fixed effects on long run corporate tax avoidance. Following Dyreng et al. (2008), we measure the long run corporate tax avoidance with a five-year CETR, i.e., CETR5.[3]

Our basic geographic unit of analysis is the MSA. These areas have high levels of sociological and economic integration. As many regional characteristics are typically measured at the MSA level, it is a good unit to use as a proxy for local community characteristics. We use the MSA-State combination to attach a unique identifier to each MSA, and this MSA-State code is included in the regression model as a separate indicator variable for each MSA. These indicators are our test variables.

Following Parsons et al. (2018), we benchmark the size of the location fixed effects against year and industry fixed effects. With this comparison, we can articulate whether the location of a firm's headquarters is likely to contain as much information as its industry classification. In a similar approach to that of Dyreng et al. (2010), who estimate executives' fixed effects, we regress the firm's tax avoidance variables on location fixed effects after controlling for year, firm and industry fixed effects. By doing so, we control time-invariant firm (industry) characteristics through firm (industry) fixed effects, and time-specific effects on corporate tax avoidance through year fixed effects. Additionally, we also control for firm attributes, e.g., size, leverage, R&D, and state corporate income tax rates, which could plausibly be related to corporate tax avoidance. The model specification is as follows:

[3] We have calculated CETR5 using (1) the sum of cash taxes paid in years t, t+1, t+2, t+3, and t+4, divided by pre-tax book income before special items in years t, t+1, t+2, t+3, and t+4; (2) the sum of cash taxes paid in years t, t-1, t-2, t-3, and t-4, divided by pre-tax book income before special items in years t, t-1, t-2, t-3, and t-4; and (3) the sum of cash taxes paid in years t-2, t-1, t, t+1, and t+2, divided by pre-tax book income before special items in years t-2, t-1, t, t+1, and t+2. The different measures all yield similar results. When CETR5 is the dependent variable, we use control variables for a firm *i* in year *t*, following prior studies such as Guenther, Matsunaga and Williams (2017) and Davis et al. (2016). In addition, we have also calculated controls based on the average of the five-year period mentioned above and the results are consistent when using controls measured at year *t*.

$$ETR_{it}(CETR_{it}\ or\ UTB_{it}\ or\ CETR5_{it-\ \sim t}) = \alpha_0 + \Sigma_l \alpha_l LOC_l + \Sigma_t \alpha_t YEAR_t + \Sigma_i \alpha_i FIRM_i + \alpha_{str} STR_{it} + \Sigma_k \alpha_k CONTROL_{it}^k + \varepsilon_{it} \quad (1)$$

$\mathrm{ETR}_{it}$, $\mathrm{CETR}_{it}$, $\mathrm{UTB}_{it}$ and $\mathrm{CETR5}_{it-4\sim t}$ are tax avoidance measures for a firm *i* in year *t*. $\mathrm{LOC}_l$ is our main variable of interest, an indicator variable for the geographic unit, i.e., MSA, in our main regression, where we denote each specific unit (location fixed effects). If the coefficients on $\mathrm{LOC}_l$ are jointly significant, the results will support our hypothesis. $\mathrm{FIRM}_i$ is an indicator variable for each firm *i* (firm fixed effects); $\mathrm{YEAR}_t$ is an indicator variable for each year *t* (year fixed effects); and $\mathrm{STR}_{it}$ is included to control for the influence of state-level tax rates. $\mathrm{CONTROL}_{it}^k$ is a vector of control variables that captures a range of time-varying firm characteristics and performance measures. $\varepsilon_{it}$ is the error term. All variables are defined in Table 7 (Variable Definitions, Appendix).

### 4.2 Location fixed effects and corporate tax avoidance

Table 3 (MSA Fixed Effects on Tax Avoidance, Appendix) reports the estimates from fixed effect regressions predicting ETRs and unrecognised tax benefits. The independent variables include year, firm, industry, and MSA fixed effects. Panels A, B, C, and D present the results when using ETR, CETR, UTB, and CETR5 as the dependent variable, respectively. We examine the F-statistics to test the joint significance of particular sets of fixed effect coefficients.

Model 1 is a baseline regression that includes only the vector of time-varying firm-level controls, STR, and an intercept. Models 2-5 include only one set of fixed effects, the effects of time-varying firm-level controls, and STR. Moving to the right in Table 3, year fixed effects are reported in Model 2, firm fixed effects in Model 3, industry fixed effects in Model 4, and MSA fixed effects in Model 5. The final model, i.e., Model 6, estimates MSA fixed effects after controlling for year fixed effects, firm fixed effects,[4] and all the time-varying firm-level variables. The results for Model 5 show that MSA fixed effects are jointly significant in explaining all of our four measures of tax avoidance (F-statistic = 1.81, 2.22, 12.68, and 5.69 for the ETR, CETR, UTB, and CETR5, respectively). These effects remain significant at less than the 1% level in Model 6, where other fixed effects are also controlled (F-statistic = 2.48, 1.95, 3.22, and 4.85 for the ETR, CETR, UTB, and CETR5 regressions, respectively). The results indicate that the MSAs in which firms are located are jointly significant in explaining the variations in corporate tax avoidance.

The explanatory power of Model 5 is benchmarked against that of the other models. In Panel A, where ETR is the dependent variable, relative to Model 1, the inclusion of year fixed effects in Model 2 increases the R-squared by roughly 0.5 percentage points, the inclusion of industry fixed effects in Model 4 increases the R-squared by about 3.6 percentage points, and the inclusion of MSA fixed effects in Model 5 increases the R-

[4] Industry fixed effects are subsumed by firm fixed effects.

squared by approximately 1.2 percentage points. The sizes of the incremental explanatory power from MSA fixed effects are similar for CETR and CETR5 as shown in Panels B and D. The impact of MSA fixed effects on the R-squared is consistently larger than that of year fixed effects, which could mean that regional variation is more important than the average differences observed across time in determining firms' ETRs. The improvements in R-squared are smaller with the MSA fixed effects than the industry fixed effects, suggesting that the location of a firm is less powerful than its industry classification in predicting how much tax the firm will avoid. However, as shown in Panel C, where UTB is the dependent variable, the R-squared increases more with MSA fixed effects than industry fixed effects. Firm fixed effects produce the largest improvement in the adjusted R-squared, as expected. However, the results from the inclusion of MSA, year and firm fixed effects in Model 6 suggest that those fixed effects are likely to capture distinctive effects on tax avoidance.

To specifically look at the explanatory power of the MSA fixed effects, we now summarise the R-squared of each model. When MSA fixed effects are added to the regressions, in comparison with the baseline model, i.e., Model 1, the R-squared increases from 7.1% to 8.3% for ETR, from 7.5% to 9.0% for CETR, from 9.4% to 18.1% for UTB, and from 16.9% to 20.5% for CETR5. This evidence is consistent with our prediction that different firm locations would at least partially account for the variations in both the one-year and long-run ETRs. It is interesting to note that including MSA fixed effects doubles the explanatory power of the UTB regression, suggesting that firm location is an important determinant of tax sheltering activity.[5]

With regard to the estimated coefficients on the firm-level control variables, most of them are statistically significant and consistent with prior studies (e.g., Gupta & Newberry, 1997; Chen et al., 2010; Dyreng et al., 2010). For example, when ETR is the dependent variable (Panel A of Table 3), the coefficients on SIZE and ROA are both positively significant at the 1% level across almost all models (except for Model 3), and this finding is consistent with larger and more profitable firms reporting more tax expenses. The estimated coefficients for ΔNOL, LEV, FI, GPPE, EQINC, MTB, R&D, CASH, ΔSALES, and ANALYST are all negative and significant, and the estimated coefficients for CAPEX, INSTPERC, SG&A and ALOCAL are positive and significant. Interestingly, the estimated coefficients on $STR_{it}$ are not statistically significant in many of the models presented in Table 3. In particular, they are −0.061, −0.001, 0.000, and 0.000 in Model 6 for the ETR, CETR, UTB, and CETR5 regressions respectively. These results show that the corporate tax avoidance practices of the sample firms are not significantly associated with the levels of state corporate income tax rates imposed on those firms, after controlling for the effects of firm, year, and location, and this finding is consistent with Hasan et al. (2017a).

Having found statistically significant location fixed effects, we next examine the economic significance of the effects of firms' locations on tax avoidance. The focus is on the MSAs whose specific fixed effect coefficient is significant, i.e., where the t-

[5] For robustness checks, we also cluster standard errors at the firm level, and then at the MSA level. We find clustering the standard errors does not change the coefficients on the fixed effects and controls variables, though has an impact on the t-statistics. In untabulated results, the F-statistics testing the joint significance of the location fixed effects are consistently significant across all regressions in the full models. In fact, in all cases, the F-statistics are slightly larger than those reported in Table 3. This suggests that our main results are robust to using clustered standard errors.

statistic on the particular coefficient is significant based on two-sided tests. Table 4 (Significance Levels of MSA Fixed Effects, Appendix) presents the numbers and percentages of individual MSAs with statistically significant fixed effects from Model 6 in Table 3 (for ETR, CETR, UTB, and CETR5 respectively). The primary sample contains a total of 142 MSA specific estimations of coefficients for which data are available.[6] In the results with ETR as the dependent variable, 49.3% of MSAs are significant at the 5% level and 35.9% do not have significant MSA fixed effect coefficients at the conventional level. The numbers of significant MSAs are larger for the CETR (93.7%) regression. For the UTB and CETR5 regressions, the percentage of location fixed effects that are significant at the 5% level are 48.1% and 62% respectively. The large numbers and percentages of significant MSAs suggest that the joint significance of the MSA fixed effect is unlikely to be driven by the influence of a few MSAs with significant coefficients.

Corresponding to this statistical significance, our results for the distribution of the MSA fixed effect coefficients estimated for ETR, CETR, UTB, and CETR5 also suggest the economic significance of the location fixed effect. The ETRs of firms located in the top 25% of MSAs are at least 13% higher than the ETRs of firms located in the bottom 25% of MSAs. This means that moving between the top and bottom quartiles of locations results in an approximately 13% swing in ETR. Given that the mean ETR is 31.5%, this difference is economically large and significant. We observe similarly sizable swings for CETR, UTB, and CETR5. We present these results in Table 6 Panel A (MSA Fixed Effects and MSA Characteristics, Appendix) for the purpose of collectively reporting the descriptive statistics for the second-stage examination. A visual depiction of the distribution of location fixed effects across the US is also provided in Figure 1 (Spatial Distribution of MSA Fixed Effect Coefficients, Appendix).

### 4.3 Robustness tests

We conduct a number of robustness tests on our main tests. The details of these tests are documented in the Supplementary Appendix. In the first set of tests, we include loss observations in our sample following recent studies such as Henry and Sansing (2018), and examine the location fixed effects on the tax avoidance behaviour of both profitable and loss firms. Results are consistent with our main findings which confirm that location fixed effects can be observed in a larger sample of firms. Second, we employ alternative measures of geographic unit, including states, counties, and the zip codes of the firms' headquarters, in estimating Equation (1). The significant location fixed effects are present when quantified using different geographic units.

Third, to address the selection bias concern given firms' change of location is not random, we split the primary sample into firms that have and firms that have not changed headquarters MSA locations and re-estimate Equation (1) using both sub-samples. Our findings that MSA fixed effects continue to be jointly significant in explaining tax avoidance activities suggest that location fixed effects are less likely to be affected by firms' decision on whether or not they change their headquarters locations. Fourth, we show that neighbouring MSAs which should have similar

[6] The number of estimated MSA fixed effects is smaller in Model 6 (142) than Model 5 (218). This is due to the inclusion of firm fixed effects in Model 6. If no firm in an MSA changes its headquarters' location during the sample period, the corresponding MSA fixed effect is subsumed by the firm fixed effects. Thus, estimated MSA fixed effects capture the location effects on firms that have moved their headquarters.

geographical features tend to influence firms' tax decisions in a similar way. Results also indicate that there is no clear pattern of location fixed effects dominating in one period of time.

The fifth set of tests address omitted variable issues. We control for corporate governance factors, such as board size, percentage of independent directors, gender diversity on the board, CEO and chair duality, and CFOs' board membership and CEO total compensation (Armstrong et al., 2015; Gaertner, 2014). The location fixed effects remain significant after we include in Equation (1) executive fixed effects (Dyreng et al., 2010; Yonker, 2017).

### 4.4 Cross-sectional tests

We expect location fixed effects to be stronger for firms with lengthier durations in their current location and lower geographic diversification. We perform two cross-sectional tests by partitioning the sample based on how long the firm has been in a given location and whether it has material subsidiaries in another US state and/or foreign country. Table 5 (MSA Fixed Effects Variation: Length of Localisation and Geographical Diversification, Appendix) presents the results.

First, we test location fixed effects on corporate tax avoidance when the firm has been located in the same MSA for more than three years (*Long*) or less than or equal to three years (*Short*). As shown in Panel A of Table 5, we find that the location fixed effects are highly significant for the *Long* sample and less or even not significant for the *Short* sample, which is consistent with location fixed effects being affected by the length of time a firm has been located in a given location.

We conjecture that the location fixed effects could attenuate for geographically diversified firms with a material presence in locations other than their headquarters locations. This is so because the location of those diversified firms' subsidiaries could dilute the location fixed effects from the headquarters. We examine this possibility using data collected by Scott D Dyreng from Exhibit 21 in 10K filings for material subsidiary disclosures.[7] The main samples are divided into firms that have material subsidiaries in at least one state or country other than the headquarter location in a given year (*Diversified*) and those that do not (*Non-diversified*). Results are reported in Panel B, Table 5. Comparing the F-statistics for MSA fixed effects across the two sub-samples suggests that in the ETR and CETR models, the location fixed effects indeed attenuate when there is geographic diversification. In the UTB regression, we find that the location fixed effects are concentrated in the geographically dispersed subsample. It is possible that, since geographically diversified firms are exposed to a variety of tax avoidance strategies (e.g., using tax haven subsidiaries to engage in aggressive tax avoidance strategies: Dyreng, Hanlon & Maydew, 2019), these firms are likely to exhibit greater variation in tax avoidance, especially in more aggressive and uncertain tax arrangements. In the CETR5 regression, the location fixed effects are also stronger in the geographically diversified subsample. This somehow contradicts our conjecture but there is a possibility that CETR5 may not reflect the geographical diversification in the current year because CETR5 represents relatively dynamic tax aggressiveness positions incorporating the past four years.

---

[7] See https://sites.google.com/site/scottdyreng/Home/data-and-code/EX21-Dataset.

## 5. MSA CHARACTERISTICS AND MSA FIXED EFFECTS

So far, we have documented significant spatial variations in corporate tax avoidance. To further understand the geographic effects on tax avoidance behaviour, we examine some possible channels through which geographic locations might affect tax avoidance decisions. Accordingly, we obtain demographic, social, economic, regulatory, and weather information about each of the MSAs in the sample, and examine whether those MSA-level characteristics are associated with the MSA fixed effects on tax avoidance.

In order to determine the extent to which MSA characteristics explain the variation in tax avoidance across MSAs, we use the following multivariate OLS regression:

$$\hat{\alpha}_l^{ETR}(\hat{\alpha}_l^{CETR}/\hat{\alpha}_l^{UTB}/\hat{\alpha}_l^{CETR5}) = \beta_0 + \beta_1 LPOP_l + \beta_2 LWAGES_l + \beta_3 LGDP_l + \beta_4 EDUC_l + \beta_5 CRIMES_l + \beta_6 RELIGION_l + \beta_7 LAUDITOR_l + \beta_8 LDISTANCEIRS_l + \beta_9 LWEATHER_l + \beta_{10} LOSTFIRMPCT_l + \epsilon_l \quad (2)$$

The dependent variables $\hat{\alpha}_l^{ETR}$, $\hat{\alpha}_l^{CETR}$, $\hat{\alpha}_l^{UTB}$, and $\hat{\alpha}_l^{CETR5}$ represent coefficients estimated for MSA $l$'s fixed effect when the dependent variables in Equation (1) are ETR, CETR, UTB, and CETR5 respectively. Recall that these coefficients are estimated while controlling for year and firm fixed effects. The independent variables include the workforce population, average personal wage of the workforce population, GDP per capita, education level of the workforce population, total crime rate, the importance of religion in people's daily life (religiosity), the number of Big 4 audit firms' local offices in the MSA, geodesic distance to IRS, and daily sunlight, for the MSAs during the sample period. Based on our discussion in section 2, we loosely group these variables into three categories: information and resource-based factors (workforce population, educational attainment, expertise from Big 4 audit firms, and proximity to IRS), factors influencing location-based risks (average personal wage, GDP per capita, and daily sunlight), and social factors (crime rates and religiosity). We also include the percentage of firms that report a negative pre-tax book income out of all firms in each MSA to proxy for the general financial performance pertained to the local area. All variables are averaged across the sample period for each MSA. Definitions and sources are provided in Panel B of Table 7 (Variable Definitions, Appendix). Table 6 (MSA Fixed Effects and MSA Characteristics, Appendix) presents the descriptive statistics for MSA characteristics, correlation between those characteristics and the relation between MSA fixed effects and MSA characteristics.

Panel A reports descriptive statistics on the regional characteristics of the MSAs in the sample for which data are available. For the 120 MSAs with available data, the average size of the workforce population is about 777,000.[8] The mean GDP per capita is USD

[8] We describe workforce population in its raw value here for a meaningful economic interpretation of the variable, but we use the log transformation of this variable in our regressions. The same strategies have

46,925 and the mean annual personal income USD 41,814. The weighted average education level of workforce population in our sample swings from 7.26 in the bottom percentile to 7.78 in the top percentile. The median average crime rate is 3.45% of the population. In general, 65.3% of Gallup poll respondents indicate that religion is an important part of their daily life. The number of Big 4 auditor offices located in the MSAs range from 0 to 13. The sample firms are located on average approximately 93 miles from the nearest IRS territory manager's office. The average daily sunlight observed across the MSAs is 16,355 KJ/ $\mathrm{m}^2$. About 6.7% of sample firms located in the MSAs report negative pre-tax book income in their financial statements.

Panel B presents the Pearson correlations for the variables in Equation (2). Overall, there appears to be some correlation between the MSA fixed effects and the location-based factors, but the correlation does not appear to be consistent across all four measures of tax avoidance.

Panel C shows the results of regressions of the MSA fixed effect coefficients on variables that capture the MSA attributes. We find the effects that firm locations have on tax avoidance to be unrelated to most of these location-based characteristics. Statistically significant coefficients are only found for the education level of the local workforce, and local sunshine coverage variables. The positive and significant coefficient estimated for $EDUC$ on the CETR and the UTB fixed effects suggests that firms located in MSAs with higher education level pay more taxes to the tax authorities and have lower tax reserves. This result is consistent with the risk explanation which suggests educated employees understand the tax risk better and are more likely comply with the tax code. Moreover, the significantly positive (negative) relation between the weather variable and the ETRs (UTB) supports the idea of a weather-induced positive mood is associated with a reduction in the risky tax avoidance activities of the firms (Chen et al., 2019).

The fact that the coefficients on most of the other independent variables in Equation (2) are insignificant suggests that, while there are significant geographic differences in tax avoidance among the MSAs in our sample, many previously examined location-related factors cannot explain these differences. The F-statistics that test the joint significance of all coefficients in each regression report a 5% significance for UTB fixed effects. The R-squareds of these models range between 7.6% (CETR5) and 16.8% (UTB). The relatively low explanatory power of all regressions indicates that the MSA characteristics studied do not account for much of the regional variation in tax avoidance observed across our sample of MSAs.

Overall, we find some evidence that location-based information and resource factors as well as risk factors are associated with persistent spatial variation in corporate tax avoidance, while local social and cultural characteristics exhibit little explanatory power. Nevertheless, we acknowledge that it is very difficult to fully identify and determine other factors that could actually account for the cross-sectional variation in tax avoidance observed across space.

---

been adopted for GDP per capita, annual personal income, the presence of Big 4 audit office, distance to IRS, and average daily sunlight.

## 6. Conclusion

This study examines whether firm location matters for tax avoidance decisions. Our findings suggest that firm location plays a significant role in predicting how much tax firms avoid, and those findings are robust across our validity tests, alternative measures tests, endogeneity tests and consistent with our cross-sectional analysis. In addition, we show that, while some observable location factors, such as education, and weather, are correlated with differences in tax avoidance across MSAs to a certain extent, much of this variation remains unexplained. Since the MSA characteristics that we examine constitute only the observable factors, there may be a set of unobservable factors that are relevant to tax avoidance decisions and tax planning activities but not captured in our models. This suggests that future studies are needed to further explore factors associated with the geographic differences in corporate tax avoidance.

Given the difficulty of identifying factors that can fully account for the substantial variation in tax avoidance across firms, documenting that firm location has a significant effect on tax avoidance and estimating the magnitude of that effect is an important step forward. Our research has implications for tax authorities, in deploying their constrained resources and coordinating their supervision efforts. The finding that firms located in some areas are more likely than others to pay lower taxes should help inform the regional enforcement of the IRS in those areas where more firms are using complex tax strategies to reduce taxes.

## 7. References

Abdul Wahab, N S & Holland, K 2012, 'Tax planning, corporate governance and equity value', *The British Accounting Review*, vol. 44, no. 2, pp. 111-124.

Arena, M P & Dewally, M 2012, 'Firm location and corporate debt', *Journal of Banking and Finance*, vol. 36, no. 4, pp. 1079-1092.

Armstrong, C S, Blouin, J L & Larcker, D F 2012, 'The incentives for tax planning', *Journal of Accounting and Economics*, vol. 53, nos 1-2, pp. 391-411.

Armstrong, C S, Blouin, J L, Jagolinzer, A D & Larcker, D F 2015, 'Corporate governance, incentives, and tax avoidance', *Journal of Accounting and Economics*, vol. 60, no. 1, pp. 1-17.

Audretsch, D B & Feldman, M P 1996, 'R&D spillovers and the geography of innovation and production', *The American Economic Review*, vol. 86, no. 3, pp. 630-640.

Audretsch, D B & Stephan, P E 1996, 'Company-scientist locational links: The case of biotechnology', *The American Economic Review*, vol. 86, no. 3, pp. 641-652.

Baik, B, Kang, J-K & Kim, J-M 2010, 'Local institutional investors, information asymmetries, and equity returns', *Journal of Financial Economics*, vol. 97, no. 1, pp. 81-106.

Baller, R D, Anselin, L, Messner, S F, Deane, G & Hawkins, D F 2001, 'Structural covariates of US county homicide rates: Incorporating spatial effects', *Criminology*, vol. 39, no. 3, pp. 561-588.

Bushee, B J 1998, 'The influence of institutional investors on myopic R&D investment behavior', *The Accounting Review*, vol. 73, no. 3, pp. 305-333.

Call, A C, Campbell, J L, Dhaliwal, D S & Moon, J R, Jr 2017, 'Employee quality and financial reporting outcomes', *Journal of Accounting and Economics*, vol. 64, no. 1, pp. 123-149.

Chen, H, Liu, S, Wang, J & Wu, Z 2022, 'The effect of geographic proximity on corporate tax avoidance: Evidence from China', *Journal of Corporate Finance* 72, 102131.

Chen, S, Chen, X, Cheng, Q & Shevlin, T 2010, 'Are family firms more tax aggressive than non-family firms?', *Journal of Financial Economics*, vol. 95, no. 1, pp. 41-61.

Chen, Y, Ge, R, Pittman, J, Veeraraghavan, M & Zolotoy, L 2019, 'Obscured by clouds: The impact of weather-induced managerial mood on corporate tax avoidance', paper presented at the 10th International Conference of the Japanese Accounting Review, Kobe, 22 December, https://www.rieb.kobe-u.ac.jp/tjar/conference/10th/.

Choi, J-H, Kim, J-B, Qiu, A A & Zang, Y 2012, 'Geographic proximity between auditor and client: How does it impact audit quality?', *Auditing: A Journal of Practice and Theory*, vol. 31, no. 2, pp. 43-72.

Clausing, K A 2009, 'Multinational firm tax avoidance and tax policy', *National Tax Journal*, vol. 62, no. 4, pp. 703-725.

Davis, A K, Guenther, D A, Krull, L K & Williams, B M 2016, 'Do socially responsible firms pay more taxes?', *The Accounting Review*, vol. 91, no. 1, pp. 47-68.

Desai, M A & Dharmapala, D 2006, 'Corporate tax avoidance and high-powered incentives', *Journal of Financial Economics*, vol. 79, no. 1, pp. 145-179.

De Simone, L, Nickerson, J, Seidman, J & Stomberg, B 2020, 'How reliably do empirical tests identify tax avoidance?', *Contemporary Accounting Research*, vol. 37, no. 3, pp. 1536-1561.

Doukas, J A & Pantzalis, C 2003, 'Geographic diversification and agency costs of debt of multinational firms', *Journal of Corporate Finance*, vol. 9, no. 1, pp. 59-92.

Dyreng, S D, Hanlon, M & Maydew, E L 2008, 'Long-run corporate tax avoidance', *The Accounting Review*, vol. 83, no. 1, pp. 61-82.

Dyreng, S D, Hanlon, M & Maydew, E L 2010, 'The effects of executives on corporate tax avoidance', *The Accounting Review*, vol. 85, no. 4, pp. 1163-1189.

Dyreng, S D, Hanlon, M & Maydew, E L 2019, 'When does tax avoidance result in tax uncertainty?', *The Accounting Review*, vol. 94, no. 2, pp. 179-203.

Gaertner, F B 2014, 'CEO after-tax compensation incentives and corporate tax avoidance', *Contemporary Accounting Research*, vol. 31, no. 4, pp. 1077-1102.

Gallemore, J, Maydew, E L & Thornock, J R 2014, 'The reputational costs of tax avoidance', *Contemporary Accounting Research*, vol. 31, no. 4, pp. 1103-1133.

Glaeser, E L, Sacerdote, B & Scheinkman, J A 1996, 'Crime and social interactions', *The Quarterly Journal of Economics*, vol. 111, no. 2, pp. 507-548.

Goetzmann, W N, Kim, D, Kumar, A & Wang, Q 2015, 'Weather-induced mood, institutional investors, and stock returns', *Review of Financial Studies*, vol. 28, no. 1, pp. 73-111.

Goetzmann, W N & Zhu, N 2005, 'Rain or shine: Where is the weather effect?', *European Financial Management*, vol. 11, no. 5, pp. 559-578.

Guenther, D A, Matsunaga, S R & Williams, B M 2017, 'Is tax avoidance related to firm risk?', *The Accounting Review*, vol. 92, no. 1, pp. 115-136.

Gupta, S & Newberry, K 1997, 'Determinants of the variability in corporate effective tax rates: Evidence from longitudinal data', *Journal of Accounting and Public Policy*, vol. 16, no. 1, pp. 1-34.

Hanlon, M & Heitzman, S 2010, 'A review of tax research', *Journal of Accounting and Economics*, vol. 50, nos 2-3, pp. 127-178.

Hasan, I, Hoi, C-K (S), Wu, Q & Zhang, H 2017a, 'Does social capital matter in corporate decisions? Evidence from corporate tax avoidance', *Journal of Accounting Research*, vol. 55, no. 3, pp. 629-668.

Hasan, M M, Al-Hadi, A, Taylor, G & Richardson, G 2017b, 'Does a firm's life cycle explain its propensity to engage in corporate tax avoidance?', *European Accounting Review*, vol. 26, no. 3, pp. 469-501.

Henry, E & Sansing, R 2018, 'Corporate tax avoidance: Data truncation and loss firms', *Review of Accounting Studies*, vol. 23, no. 3, pp. 1042-1070.

Hjelström, T, Kallunki, J-P, Nilsson, H & Tylaite, M 2020, 'Executives' personal tax behavior and corporate tax avoidance consistency', *European Accounting Review* vol. 29, no. 3, pp. 493-520.

Hoopes, J L, Mescall, D & Pittman, J A 2012, 'Do IRS audits deter corporate tax avoidance?', *The Accounting Review*, vol. 87, no. 5, pp. 1603-1639.

Jaffe, A B, Trajtenberg, M & Henderson, R 1993, 'Geographic localization of knowledge spillovers as evidenced by patent citations', *The Quarterly Journal of Economics*, vol. 108, no. 3, pp. 577-598.

John, K, Knyazeva, A & Knyazeva, D 2011, 'Does geography matter? Firm location and corporate payout policy', *Journal of Financial Economics*, vol. 101, no. 3, pp. 533-551.

Jung, S 2015, 'Does education affect risk aversion? Evidence from the British education reform', *Applied Economics*, vol. 47, no. 28, pp. 2924-2938.

Kamstra, M J, Kramer, L A & Levi, M D 2003, 'Winter blues: A SAD stock market cycle', *American Economic Review*, vol. 93, no. 1, pp. 324-343.

Kedia, S & Rajgopal, S 2011, 'Do the SEC's enforcement preferences affect corporate misconduct?', *Journal of Accounting and Economics*, vol. 51, no. 3, pp. 259-278.

Kubick, T R, Lockhart, G B, Mills, L F & Robinson, J R 2017, 'IRS and corporate taxpayer effects of geographic proximity', *Journal of Accounting and Economics*, vol. 63, nos 2-3, pp. 428-453.

Land, K C, McCall, P L & Cohen, L E 1990, 'Structural covariates of homicide rates: Are there any invariances across time and social space?', *American Journal of Sociology*, vol. 95, no. 4, pp. 922-963.

Lisowsky, P, Robinson, L & Schmidt, A 2013, 'Do publicly disclosed tax reserves tell us about privately disclosed tax shelter activity?', *Journal of Accounting Research*, vol. 51, no. 3, pp. 583-629.

Malloy, C J 2005, 'The geography of equity analysis', *The Journal of Finance*, vol. 60, no. 2, pp. 719-755.

McGuire, S T, Omer, T C & Sharp, N Y 2012a, 'The impact of religion on financial reporting irregularities', *The Accounting Review*, vol. 87, no. 2, pp. 645-673.

McGuire, S T, Omer, T C & Wang, D 2012b, 'Tax avoidance: Does tax-specific industry expertise make a difference?', *The Accounting Review*, vol. 87, no. 3, pp. 975-1003.

Mills, L F 1998, 'Book-tax differences and Internal Revenue Service adjustments', *Journal of Accounting Research*, vol. 36, no. 2, pp. 343-356.

Myers, J N, Myers, L A & Skinner, D J 2007, 'Earnings momentum and earnings management', *Journal of Accounting, Auditing and Finance*, vol. 22, no. 2, pp. 249-284.

Parsons, C A, Sulaeman, J & Titman, S 2018, 'The geography of financial misconduct', *The Journal of Finance*, vol. 73, no. 5, pp. 2087-2137.

Rego, S O 2003, 'Tax-avoidance activities of US multinational corporations', *Contemporary Accounting Research*, vol. 20, no. 4, pp. 805-833.

Richardson, G & Lanis, R 2007, 'Determinants of the variability in corporate effective tax rates and tax reform: Evidence from Australia', *Journal of Accounting and Public Policy*, vol. 26, no. 6, pp. 689-704.

Sampson, R J, Raudenbush, S W & Earls, F 1997, 'Neighborhoods and violent crime: A multilevel study of collective efficacy', *Science*, vol. 227, no. 5328, pp. 918-924.

Shaw, C R & McKay, H D 1942, *Juvenile delinquency and urban areas: A study of rates of delinquents in relation to differential characteristics of local communities in American cities*, University of Chicago Press, Chicago.

Sherman, L W & Weisburd, D 1995, 'General deterrent effects of police patrol in crime "hot spots": A randomized, controlled trial', *Justice Quarterly*, vol. 12, no. 4, pp. 625-648.

Skolnick, J H & Bayley, D H 1986, *The new blue line: Police innovation in six American cities*, Free Press, New York.

Su, K, Li, B & Ma, C 2019, 'Corporate dispersion and tax avoidance', *Chinese Management Studies*, vol. 13, no. 3, pp. 706-732.

Weisburd, D 2015, 'The law of crime concentration and the criminology of place', *Criminology*, vol. 53, no. 2, pp. 133-157.

Yonker, S E 2017, 'Geography and the market for CEOs', *Management Science*, vol. 63, no. 3, pp. 609-630.

Zimmerman, J L 1983, 'Taxes and firm size', *Journal of Accounting and Economics*, vol. 5, pp. 119-149.

## 8. APPENDIX



**Table 1: Sample Selection**

**Panel A: Sample Selection for the Primary Sample**

| | ETR/CETR | UTB | CETR5 |
|---|---|---|---|
| Firm-year observations relating firms located in the US from 1993 (or 2006 for the UTB sample) to 2017 with sufficient data on COMPUSTAT (observations in 1993 (or 2006) are included for calculating lagged values) | 220,282 | 95,271 | 220,282 |
| Less: Financial companies (SIC 6000-6999) and highly regulated industry (SIC 4400-5000) | 84,819 | 43,967 | 84,819 |
| Less: Firm-year observations with missing or invalid location data (zip code, state, county, MSA) | 39,555 | 9,909 | 39,555 |
| Less: Firm-year observations with negative pre-tax book income, negative income tax expense, or negative cash taxes paid (for CETR5, observations with negative pre-tax income when summed over the five-year period, i.e. negative denominators) | 47,849 | ---------- | 56,510 |
| Less: Firm-year observations with insufficient data to calculate tax avoidance measures ETR and CETR (or CETR5 or UTB) | 4,356 | 16,369 | 9,722 |
| Less: Firm-year observations with insufficient data to calculate control variables | 12,556 | 5,266 | 7,200 |
| Less: MSAs with less than 2 firms in each year | 1,854 | 835 | 1,769 |
| Final Sample (firm-year) | 29,293 | 18,925 | 20,707 |
| Final Sample (firm) | 5,197 | 3,322 | 3,248 |
| Final Sample (MSA) | 218 | 158 | 164 |

| | | | |
|---|---|---|---|
| Final Sample (state) | 50 | 44 | 47 |
| Final Sample (county) | 557 | 375 | 419 |
| Final Sample (zip code) | 2,769 | 1,814 | 2,046 |

**Panel B: Sample Selection for the MSA Sample**

| | ETR/CETR | UTB | CETR5 |
|---|---|---|---|
| MSAs with coefficients estimated in Equation (1) | 218 | 158 | 164 |
| Less: MSAs with missing fixed effect coefficients (MSA fixed effects subsumed by firm fixed effects) | 76 | 54 | 72 |
| Less: MSAs with missing values in Equation (2) variables | 22 | 18 | 15 |
| Final sample MSAs | 120 | 86 | 77 |

**Table 2: Descriptive Statistics**

| | N | Mean | Std. Dev. | $25^{th}$ Pctl | Median | $75^{th}$ Pctl |
|---|---|---|---|---|---|---|
| **Dependent Variables** | | | | | | |
| $\mathrm{ETR}_{it}$ | 29,293 | 0.315 | 0.174 | 0.241 | 0.335 | 0.381 |
| $\mathrm{CETR}_{it}$ | 29,293 | 0.259 | 0.227 | 0.102 | 0.241 | 0.349 |
| $\mathrm{UTB}_{it}$ | 18,925 | 0.012 | 0.023 | 0.000 | 0.004 | 0.013 |
| $\mathrm{CETR5}_{it-4\sim t}$ | 20,707 | 0.256 | 0.119 | 0.180 | 0.268 | 0.338 |
| **Control Variables** | | | | | | |
| $\mathrm{SIZE}_{it}$ | 29,293 | 5.988 | 2.004 | 4.616 | 6.033 | 7.364 |
| $\mathrm{ROA}_{it}$ | 29,293 | 0.129 | 0.107 | 0.057 | 0.102 | 0.169 |
| $\mathrm{NOL}_{it}$ | 29,293 | 0.412 | 0.492 | 0.000 | 0.000 | 1.000 |
| $\Delta\,\mathrm{NOL}_{it}$ | 29,293 | 0.111 | 0.390 | 0.000 | 0.000 | 0.040 |
| $\mathrm{LEV}_{it}$ | 29,293 | 0.201 | 0.193 | 0.015 | 0.168 | 0.320 |
| $\mathrm{FI}_{it}$ | 29,293 | 0.018 | 0.037 | 0.000 | 0.000 | 0.021 |
| $\mathrm{GPPE}_{it}$ | 29,293 | 0.489 | 0.366 | 0.200 | 0.395 | 0.699 |
| $\mathrm{INTANG}_{it}$ | 29,293 | 0.164 | 0.182 | 0.010 | 0.098 | 0.267 |
| $\mathrm{EQINC}_{it}$ | 29,293 | 0.001 | 0.004 | 0.000 | 0.000 | 0.000 |
| $\mathrm{MTB}_{it}$ | 29,293 | 2.969 | 3.379 | 1.331 | 2.141 | 3.533 |
| $\mathrm{R\&D}_{it}$ | 29,293 | 0.032 | 0.059 | 0.000 | 0.000 | 0.034 |
| $\mathrm{CASH}_{it}$ | 29,293 | 0.162 | 0.177 | 0.027 | 0.095 | 0.240 |
| $\mathrm{CAPEX}_{it}$ | 29,293 | 0.128 | 0.097 | 0.062 | 0.100 | 0.163 |
| $\mathrm{ADV}_{it}$ | 29,293 | 0.011 | 0.026 | 0.000 | 0.000 | 0.008 |
| $\mathrm{SG\&A}_{it}$ | 29,293 | 0.239 | 0.167 | 0.112 | 0.212 | 0.330 |
| $\Delta\,\mathrm{SALES}_{it}$ | 29,293 | 0.154 | 0.284 | 0.013 | 0.093 | 0.215 |
| $\mathrm{INSTPERC}_{it}$ | 29,293 | 0.575 | 0.328 | 0.413 | 0.695 | 0.823 |
| $\mathrm{MKTPRES}_{it}$ | 29,293 | 1.947 | 4.777 | 0.000 | 0.000 | 2.000 |
| $\mathrm{ANALYST}_{it}$ | 29,293 | 1.796 | 1.470 | 0.000 | 2.079 | 3.045 |
| $\mathrm{ALOCAL}_{it}$ | 29,293 | 0.594 | 0.590 | 0.000 | 1.000 | 1.000 |
| $\mathrm{STR}_{it}$ | 29,293 | 6.774 | 3.032 | 6.000 | 7.500 | 8.840 |

This Table reports descriptive statistics for the variables used in our main analyses. All continuous tax and control variables (except for STR) are winsorised at the 1% and 99% level to mitigate the influence of outliers. All variables are defined in Panel A of Table 7.

**Table 3: MSA Fixed Effects on Tax Avoidance**

**Panel A: MSA Fixed Effects on Effective Tax Rates ($ETR_{it}$)**

| | **Model 1** | **Model 2** | **Model 3** | **Model 4** | **Model 5** | **Model 6** |
|---|---|---|---|---|---|---|
| Joint Significance (F-statistics) | | | | | | |
| $YEAR_t$ | | 7.23*** | | | | 7.56*** |
| $FIRM_i$ | | | 3.01*** | | | 3.05*** |
| $IND_i$ | | | | 3.20*** | | ---------- |
| $LOC_l$ (MSA) | | | | | 1.81*** | 2.48*** |
| N | 29,293 | 29,293 | 29,293 | 29,293 | 29,293 | 29,293 |
| NYEARS | | 24 | | | | 24 |
| NFIRMS | | | 5,197 | | | 5,197 |
| NINDS | | | | 367 | | ---------- |
| NLOCS | | | | | 218 | 218 |
| RSQ | 0.071 | 0.076 | 0.437 | 0.107 | 0.083 | 0.449 |
| Estimated Coefficients (t-tests): | | | | | | |
| Intercept | 0.313*** | 0.340*** | 0.208*** | 0.160 | 0.352*** | −0.733*** |
| | (48.71) | (38.91) | (2.76) | (1.37) | (7.98) | (−3.35) |
| $SIZE_{it}$ | 0.005*** | 0.006*** | 0.004* | 0.005*** | 0.005*** | 0.012*** |
| | (6.29) | (7.05) | (1.69) | (5.81) | (6.18) | (4.43) |
| $ROA_{it}$ | 0.108*** | 0.107*** | 0.204*** | 0.114*** | 0.120*** | 0.205*** |
| | (9.82) | (9.66) | (14.36) | (10.17) | (9.83) | (14.44) |
| $NOL_{it}$ | −0.002 | 0.002 | −0.006* | −0.001 | −0.001 | −0.002 |
| | (−0.80) | (0.99) | (−1.90) | (−0.53) | (−0.44) | (−0.65) |
| $\Delta$ $NOL_{it}$ | −0.065*** | −0.064*** | −0.059*** | −0.067*** | −0.067*** | −0.054*** |
| | (−22.83) | (−22.27) | (−12.33) | (−23.22) | (−23.18) | (−11.00) |
| $LEV_{it}$ | −0.036*** | −0.042*** | 0.028*** | −0.031*** | −0.035*** | 0.011 |
| | (−6.00) | (−6.91) | (2.90) | (−4.85) | (−5.83) | (1.10) |
| $FI_{it}$ | −0.369*** | −0.365*** | −0.325*** | −0.299*** | −0.373*** | −0.307*** |
| | (−12.66) | (−12.53) | (−7.01) | (−9.58) | (−12.53) | (−6.55) |
| $GPPE_{it}$ | −0.033*** | −0.031*** | −0.038*** | −0.038*** | −0.031*** | −0.026*** |
| | (−9.68) | (−9.09) | (−4.30) | (−8.30) | (−8.80) | (−2.90) |
| $INTANG_{it}$ | −0.013* | −0.002 | −0.052*** | −0.030*** | −0.014** | −0.045*** |
| | (−1.85) | (−0.36) | (−3.67) | (−3.78) | (−1.95) | (−3.16) |
| $EQINC_{it}$ | −0.813*** | −0.809*** | −0.227 | −0.802*** | −0.801*** | −0.230 |
| | (−3.58) | (−3.57) | (−0.69) | (−3.35) | (−3.48) | (−0.70) |
| $MTB_{it}$ | −0.001*** | −0.001*** | −0.002*** | −0.002*** | −0.001*** | −0.002*** |
| | (−4.44) | (−4.45) | (−5.01) | (−4.89) | (−4.54) | (−5.00) |
| $R\&D_{it}$ | −0.354*** | −0.363*** | −0.097 | −0.213*** | −0.347*** | −0.119* |
| | (−15.65) | (−16.07) | (−1.54) | (−7.54) | (−14.50) | (−1.87) |
| $CASH_{it}$ | −0.055*** | −0.044*** | −0.063*** | −0.064*** | −0.056*** | −0.051*** |
| | (−7.11) | (−5.50) | (−5.09) | (−7.61) | (−7.01) | (−4.03) |
| $CAPEX_{it}$ | 0.124*** | 0.112*** | 0.043*** | 0.106*** | 0.122*** | 0.027* |
| | (10.92) | (9.67) | (3.10) | (8.98) | (10.61) | (1.90) |
| $ADV_{it}$ | 0.018 | 0.031 | 0.117 | −0.022 | 0.005 | 0.108 |
| | (0.44) | (0.75) | (1.24) | (−0.47) | (0.12) | (1.14) |

| | | | | | | |
|---|---|---|---|---|---|---|
| $SG\&A_{it}$ | 0.041*** | 0.041*** | 0.038** | 0.034*** | 0.042*** | 0.058*** |
| | (5.06) | (5.04) | (2.02) | (3.67) | (5.06) | (3.07) |
| $\Delta\ SALES_{it}$ | −0.038*** | −0.040*** | −0.037*** | −0.035*** | −0.036*** | −0.039*** |
| | (−9.82) | (−10.18) | (−8.54) | (−8.90) | (−9.29) | (−8.93) |
| $INSTPERC_{it}$ | 0.020*** | 0.024*** | 0.004 | 0.021*** | 0.020*** | 0.010** |
| | (5.36) | (6.35) | (0.95) | (5.72) | (5.37) | (2.01) |
| $MKTPRES_{it}$ | −0.000 | −0.000 | −0.000 | −0.000 | −0.000 | −0.000 |
| | (−0.41) | (−0.20) | (−0.48) | (−1.49) | (−0.29) | (−0.58) |
| $ANALYST_{it}$ | −0.008*** | −0.007*** | −0.006*** | −0.008*** | −0.008*** | −0.006*** |
| | (−6.83) | (−6.23) | (−3.84) | (−7.17) | (−6.78) | (−3.45) |
| $ALOCAL_{it}$ | 0.007*** | 0.005*** | 0.001 | 0.007*** | 0.007*** | 0.003 |
| | (3.49) | (2.71) | (0.23) | (3.30) | (3.37) | (0.95) |
| $STR_{it}$ | 0.000 | −0.000 | 0.000 | 0.000 | 0.091 | −0.061 |
| | (0.87) | (−0.02) | (0.06) | (0.22) | (1.11) | (−0.69) |

**Panel B: MSA Fixed Effects on Cash Effective Tax Rates ($CETR_{it}$)**

| | **Model 1** | **Model 2** | **Model 3** | **Model 4** | **Model 5** | **Model 6** |
|---|---|---|---|---|---|---|
| Joint Significance (F-statistics) | | | | | | |
| $YEAR_t$ | | 11.59*** | | | | 12.10*** |
| $FIRM_i$ | | | 2.54*** | | | 2.54*** |
| $IND_i$ | | | | 2.98*** | | ---------- |
| $LOC_l$ (MSA) | | | | | 2.22*** | 1.95*** |
| N | 29,293 | 29,293 | 29,293 | 29,293 | 29,293 | 29,293 |
| NYEARS | | 24 | | | | 24 |
| NFIRMS | | | 5,197 | | | 5,197 |
| NINDS | | | | 367 | | ---------- |
| NLOCS | | | | | 218 | 218 |
| RSQ | 0.075 | 0.083 | 0.403 | 0.109 | 0.090 | 0.416 |
| Estimated Coefficients (t-tests): | | | | | | |
| Intercept | 0.328*** | 0.314*** | 0.120 | 0.156 | 0.523*** | −0.984*** |
| | (39.09) | (27.58) | (1.18) | (1.02) | (10.64) | (−3.34) |
| $SIZE_{it}$ | 0.004*** | 0.005*** | 0.016*** | 0.004*** | 0.004*** | 0.033*** |
| | (3.97) | (4.99) | (5.46) | (3.78) | (3.86) | (8.99) |
| $ROA_{it}$ | −0.066*** | −0.072*** | −0.141*** | −0.078*** | −0.072*** | −0.152*** |
| | (−4.61) | (−4.99) | (−7.36) | (−5.28) | (−4.92) | (−7.96) |
| $NOL_{it}$ | −0.037*** | −0.033*** | −0.032*** | −0.036*** | −0.037*** | −0.026*** |
| | (−12.70) | (−10.86) | (−8.02) | (−11.93) | (−12.34) | (−6.34) |
| $\Delta\ NOL_{it}$ | −0.055*** | −0.052*** | −0.036*** | −0.054*** | −0.057*** | −0.023*** |
| | (−14.63) | (−13.81) | (−5.62) | (−14.23) | (−15.04) | (−3.56) |
| $LEV_{it}$ | −0.086*** | −0.096*** | 0.001 | −0.073*** | −0.085*** | −0.025* |
| | (−10.97) | (−12.24) | (0.09) | (−8.79) | (−10.71) | (−1.87) |
| $FI_{it}$ | −0.123*** | −0.122*** | −0.531*** | −0.077* | −0.119*** | −0.507*** |
| | (−3.23) | (−3.20) | (−8.50) | (1.88) | (−3.07) | (−8.03) |
| $GPPE_{it}$ | −0.065*** | −0.062*** | −0.028** | −0.039*** | −0.061*** | 0.003 |

| | | | | | | |
|---|---|---|---|---|---|---|
| | (−14.77) | (−14.11) | (−2.40) | (−6.54) | (−13.39) | (0.21) |
| $\mathrm{INTANG}_{it}$ | −0.031*** | −0.015 | −0.051*** | −0.022** | −0.032*** | −0.016 |
| | (−3.44) | (−1.64) | (−2.68) | (−2.12) | (−3.53) | (−0.84) |
| $\mathrm{EQINC}_{it}$ | −1.036*** | −0.989*** | −1.233*** | −0.839*** | −0.949*** | −1.131** |
| | (−3.49) | (−3.34) | (−2.79) | (−2.68) | (−3.16) | (−2.55) |
| $\mathrm{MTB}_{it}$ | −0.001*** | −0.001*** | −0.001*** | −0.002*** | −0.001*** | −0.001** |
| | (−3.22) | (−2.99) | (−2.77) | (−3.61) | (−3.18) | (−2.10) |
| $\mathrm{R\&D}_{it}$ | −0.469*** | −0.483*** | −0.045 | −0.348*** | −0.445*** | −0.048 |
| | (−15.88) | (−16.38) | (−0.53) | (−9.43) | (−14.26) | (−0.56) |
| $\mathrm{CASH}_{it}$ | −0.062*** | −0.042*** | −0.071*** | −0.039*** | −0.055*** | −0.030* |
| | (−6.07) | (−4.10) | (−4.22) | (−3.52) | (−5.32) | (−1.74) |
| $\mathrm{CAPEX}_{it}$ | 0.051*** | 0.015 | 0.133*** | −0.088*** | 0.060*** | 0.077*** |
| | (3.44) | (1.00) | (7.06) | (5.66) | (4.01) | (4.03) |
| $\mathrm{ADV}_{it}$ | 0.046 | 0.073 | −0.083 | −0.024 | 0.035 | −0.066 |
| | (0.85) | (1.37) | (−0.65) | (−0.39) | (0.63) | (−0.52) |
| $\mathrm{SG\&A}_{it}$ | 0.072*** | 0.076*** | 0.153*** | 0.076*** | 0.073*** | 0.164*** |
| | (6.84) | (7.18) | (5.99) | (6.25) | (6.78) | (6.42) |
| $\Delta\ \mathrm{SALES}_{it}$ | −0.096*** | −0.097*** | −0.079*** | −0.086*** | −0.091*** | −0.084*** |
| | (−19.20) | (−19.12) | (−13.60) | (−16.98) | (−18.08) | (−14.14) |
| $\mathrm{INSTPERC}_{it}$ | 0.033*** | 0.039*** | 0.002 | 0.033*** | 0.031*** | 0.007 |
| | (6.92) | (7.92) | (0.38) | (6.65) | (6.52) | (1.08) |
| $\mathrm{MKTPRES}_{it}$ | −0.000 | 0.000 | −0.000 | −0.000 | −0.000 | −0.000 |
| | (−0.11) | (0.33) | (−0.57) | (−0.99) | (−0.12) | (−1.05) |
| $\mathrm{ANALYST}_{it}$ | −0.009*** | −0.008*** | −0.000 | −0.009*** | −0.008*** | −0.000 |
| | (−5.82) | (−5.70) | (−0.18) | (−6.09) | (−5.52) | (−0.19) |
| $\mathrm{ALOCAL}_{it}$ | 0.005* | 0.003 | 0.007* | 0.006** | 0.005* | 0.007* |
| | (1.95) | (1.16) | (1.85) | (2.44) | (1.96) | (1.90) |
| $\mathrm{STR}_{it}$ | 0.001** | 0.000 | 0.001 | 0.000 | 0.001 | −0.001 |
| | (2.24) | (0.92) | (1.29) | (0.87) | (0.82) | (−0.56) |

**Panel C: MSA Fixed Effects on FIN 48 Tax Reserve ($\mathbf{UTB}_{it}$)**

| | **Model 1** | **Model 2** | **Model 3** | **Model 4** | **Model 5** | **Model 6** |
|---|---|---|---|---|---|---|
| Joint Significance (F-statistics) | | | | | | |
| $\mathrm{YEAR}_t$ | | 10.65*** | | | | 4.37*** |
| $\mathrm{FIRM}_i$ | | | 16.70*** | | | 15.12*** |
| $\mathrm{IND}_i$ | | | | 5.85*** | | ---------- |
| $\mathrm{LOC}_l$ (MSA) | | | | | 12.68*** | 3.22*** |
| N | 18,925 | 18,925 | 18,925 | 18,925 | 18,925 | 18,925 |
| NYEARS | | 11 | | | | 11 |
| NFIRMS | | | 3,322 | | | 3,322 |
| NINDS | | | | 326 | | ---------- |
| NLOCS | | | | | 158 | 158 |
| RSQ | 0.094 | 0.099 | 0.801 | 0.178 | 0.181 | 0.806 |

Estimated Coefficients (t-tests):

| | | | | | | |
|---|---|---|---|---|---|---|
| Intercept | −0.001 | −0.003*** | 0.124*** | −0.007** | 0.001 | 0.112*** |
| | (−0.97) | (−2.58) | (17.68) | (−2.08) | (0.18) | (11.55) |
| $SIZE_{it}$ | 0.000*** | 0.000*** | −0.006*** | 0.001*** | 0.000*** | −0.006*** |
| | (3.67) | (3.71) | (−22.24) | (7.36) | (4.01) | (−21.36) |
| $ROA_{it}$ | 0.003*** | 0.003*** | 0.001** | 0.003*** | 0.003*** | 0.001** |
| | (6.85) | (6.86) | (2.22) | (6.09) | (6.81) | (2.54) |
| $NOL_{it}$ | 0.002*** | 0.003*** | 0.000 | 0.002*** | 0.001*** | 0.000 |
| | (5.25) | (6.86) | (0.22) | (4.75) | (3.31) | (0.23) |
| $\Delta NOL_{it}$ | 0.001*** | 0.001*** | 0.001*** | 0.001*** | 0.001*** | 0.000*** |
| | (15.03) | (15.44) | (9.55) | (14.74) | (14.04) | (8.75) |
| $LEV_{it}$ | −0.001 | 0.000 | 0.001* | 0.001 | 0.000 | 0.001* |
| | (−0.86) | (−0.68) | (1.83) | (1.01) | (−0.02) | (1.72) |
| $FI_{it}$ | 0.026*** | 0.025*** | −0.003 | 0.013*** | 0.023*** | −0.003 |
| | (6.20) | (5.99) | (−0.71) | (3.13) | (5.61) | (−0.74) |
| $GPPE_{it}$ | 0.000 | 0.001 | 0.004*** | 0.003*** | 0.001*** | 0.003*** |
| | (0.92) | (1.15) | (4.98) | (4.08) | (2.86) | (4.30) |
| $INTANG_{it}$ | 0.002 | 0.002** | 0.001 | −0.003** | 0.002* | 0.001 |
| | (1.59) | (1.98) | (0.62) | (−2.28) | (1.77) | (0.45) |
| $EQINC_{it}$ | 0.051 | 0.050 | 0.027 | 0.095** | 0.084* | 0.032 |
| | (1.14) | (1.11) | (0.67) | (2.05) | (1.92) | (0.80) |
| $MTB_{it}$ | 0.000 | 0.000** | 0.000 | 0.000*** | 0.000 | 0.000 |
| | (1.64) | (2.31) | (1.24) | (2.59) | (0.71) | (1.21) |
| $R\&D_{it}$ | −0.000** | −0.000** | 0.000 | 0.000 | 0.000 | 0.000 |
| | (−2.56) | (−2.45) | (0.59) | (0.47) | (0.25) | (1.15) |
| $CASH_{it}$ | 0.022*** | 0.022*** | −0.001 | 0.013*** | 0.013*** | −0.001 |
| | (21.84) | (21.96) | (−0.36) | (11.62) | (13.20) | (−0.86) |
| $CAPEX_{it}$ | −0.016*** | −0.016*** | −0.005*** | −0.017*** | −0.019*** | −0.004*** |
| | (−8.95) | (−8.99) | (−3.49) | (−9.21) | (−10.95) | (−2.75) |
| $ADV_{it}$ | 0.002 | 0.003 | 0.009 | 0.009 | 0.011** | 0.016** |
| | (0.31) | (0.50) | (1.23) | (1.51) | (2.12) | (2.16) |
| $SG\&A_{it}$ | 0.001*** | 0.001*** | 0.001*** | 0.001*** | 0.001*** | 0.001*** |
| | (4.94) | (4.95) | (3.77) | (4.69) | (4.56) | (3.16) |
| $\Delta SALES_{it}$ | −0.002*** | −0.002*** | −0.001*** | −0.001*** | −0.001*** | −0.001*** |
| | (−5.08) | (−4.95) | (−2.78) | (−3.57) | (−4.46) | (−2.69) |
| $INSTPERC_{it}$ | −0.001** | −0.002** | 0.000 | −0.001 | 0.001 | −0.000 |
| | (−2.10) | (−2.52) | (0.20) | (−0.91) | (0.90) | (−0.09) |
| $MKTPRES_{it}$ | −0.000*** | −0.000*** | −0.000** | −0.000*** | −0.000*** | −0.000** |
| | (−5.91) | (−5.44) | (−2.52) | (−4.62) | (−4.85) | (−2.30) |
| $ANALYST_{it}$ | −0.008* | −0.000 | −0.000 | −0.001*** | −0.001*** | −0.000 |
| | (−1.94) | (−1.02) | (−0.46) | (−3.76) | (−4.09) | (−0.35) |
| $ALOCAL_{it}$ | 0.003*** | 0.003*** | −0.000 | 0.003*** | 0.002*** | 0.000 |
| | (9.19) | (8.15) | (−0.43) | (8.00) | (6.29) | (0.30) |
| $STR_{it}$ | 0.001*** | 0.001*** | −0.000*** | 0.000*** | 0.000 | 0.000 |
| | (12.44) | (11.72) | (−2.70) | (8.41) | (0.61) | (1.57) |

| Panel D: MSA Fixed Effects on 5-Year Cash ETR ($CETR5_{it-4\sim t}$) | Model 1 | Model 2 | Model 3 | Model 4 | Model 5 | Model 6 |
|---|---|---|---|---|---|---|
| Joint Significance (F-statistics) | | | | | | |
| $YEAR_t$ | | 26.56*** | | | | 32.10*** |
| $FIRM_i$ | | | 10.01*** | | | 10.08*** |
| $IND_i$ | | | | 9.46*** | | ---------- |
| $LOC_l$ (MSA) | | | | | 5.69*** | 4.85*** |
| N | 20,707 | 20,707 | 20,707 | 20,707 | 20,707 | 20,707 |
| NYEARS | | 24 | | | | 24 |
| NFIRMS | | | 3,248 | | | 3,248 |
| NINDS | | | | 355 | | ---------- |
| NLOCS | | | | | 164 | 164 |
| RSQ | 0.169 | 0.192 | 0.710 | 0.287 | 0.205 | 0.728 |
| Estimated Coefficients (t-tests): | | | | | | |
| Intercept | 0.315*** | 0.297*** | 0.240*** | 0.211*** | 0.298*** | −0.043 |
| | (57.28) | (43.86) | (6.33) | (12.99) | (26.03) | (−0.80) |
| $SIZE_{it}$ | 0.001** | 0.003*** | −0.001 | 0.002** | 0.001* | 0.016*** |
| | (2.04) | (4.85) | (−0.79) | (2.33) | (1.65) | (8.80) |
| $ROA_{it}$ | 0.128*** | 0.125*** | 0.013 | 0.116*** | 0.126*** | 0.003 |
| | (12.98) | (12.80) | (1.43) | (11.93) | (12.74) | (0.34) |
| $NOL_{it}$ | −0.019*** | −0.015*** | −0.016*** | −0.018*** | −0.018*** | −0.013*** |
| | (−10.48) | (−7.99) | (−8.63) | (−9.90) | (−9.90) | (−6.79) |
| $\Delta NOL_{it}$ | −0.184*** | −0.177*** | −0.113*** | −0.178*** | −0.189*** | −0.096*** |
| | (−22.78) | (−22.18) | (−12.84) | (−22.88) | (−23.48) | (−11.08) |
| $LEV_{it}$ | −0.063*** | −0.079*** | 0.020*** | −0.048*** | −0.064*** | −0.006 |
| | (−12.78) | (−15.83) | (3.33) | (−9.35) | (−12.80) | (−0.97) |
| $FI_{it}$ | −0.164*** | −0.160*** | −0.264*** | −0.137*** | −0.170*** | −0.241*** |
| | (−7.15) | (−7.06) | (−9.41) | (−5.76) | (−7.33) | (−8.73) |
| $GPPE_{it}$ | −0.056*** | −0.052*** | −0.009 | −0.025*** | −0.052*** | 0.017*** |
| | (−20.14) | (−18.69) | (−1.38) | (−6.55) | (−18.20) | (2.66) |
| $INTANG_{it}$ | −0.034*** | −0.012** | −0.047*** | −0.022*** | −0.036*** | −0.015* |
| | (−6.42) | (−2.19) | (−5.32) | (−3.62) | (−6.85) | (−1.73) |
| $EQINC_{it}$ | −0.981*** | −0.899*** | −1.208*** | −0.595*** | −0.893*** | −1.035*** |
| | (−5.33) | (−4.95) | (−5.67) | (−3.15) | (−4.83) | (−4.96) |
| $MTB_{it}$ | −0.001*** | −0.001*** | −0.001*** | −0.002*** | −0.001*** | −0.000 |
| | (−4.79) | (−4.54) | (−3.57) | (−6.74) | (−5.06) | (−1.51) |
| $R\&D_{it}$ | −0.431*** | −0.451*** | −0.133*** | −0.283*** | −0.392*** | −0.105** |
| | (−22.49) | (−23.80) | (−3.00) | (−11.87) | (−19.39) | (−2.40) |
| $CASH_{it}$ | −0.070*** | −0.046*** | −0.056*** | −0.038*** | −0.065*** | −0.019** |
| | (−10.74) | (−6.92) | (−6.92) | (−5.50) | (−9.75) | (−2.33) |
| $CAPEX_{it}$ | −0.085*** | −0.110*** | 0.017* | −0.041*** | −0.073*** | −0.004 |
| | (−8.29) | (−10.69) | (1.76) | (−4.00) | (−7.10) | (−0.47) |
| $ADV_{it}$ | 0.071** | 0.099*** | −0.084 | 0.046 | 0.068** | −0.021 |

| | | | | | | |
|---|---|---|---|---|---|---|
| | (2.23) | (3.13) | (−1.44) | (1.30) | (2.08) | (−0.37) |
| $SG\&A_{it}$ | 0.077*** | 0.078*** | 0.028* | 0.056*** | 0.074*** | 0.038*** |
| | (11.21) | (11.45) | (1.95) | (6.92) | (10.40) | (2.77) |
| $\Delta$ $SALES_{it}$ | −0.065*** | −0.065*** | −0.026*** | −0.051*** | −0.061*** | −0.025*** |
| | (−14.64) | (−14.36) | (−7.49) | (−11.83) | (−13.80) | (−7.17) |
| $INSTPERC_{it}$ | 0.016*** | 0.030*** | −0.002 | 0.014*** | 0.014*** | 0.013*** |
| | (5.36) | (9.54) | (0.83) | (4.57) | (4.75) | (4.18) |
| $MKTPRES_{it}$ | 0.001*** | 0.001*** | 0.000 | 0.000** | 0.001*** | 0.000 |
| | (3.98) | (4.44) | (1.57) | (2.46) | (4.09) | (0.88) |
| $ANALYST_{it}$ | −0.008*** | −0.008*** | −0.001 | −0.009*** | −0.007*** | −0.003*** |
| | (−8.77) | (−9.51) | (−1.23) | (−10.44) | (−8.34) | (−3.84) |
| $ALOCAL_{it}$ | 0.003** | 0.002 | −0.001 | 0.006*** | 0.003** | 0.001 |
| | (2.34) | (1.01) | (−0.34) | (3.93) | (2.06) | (0.58) |
| $STR_{it}$ | 0.002*** | 0.001*** | 0.001 | 0.001*** | 0.000 | 0.000 |
| | (5.89) | (4.00) | (1.53) | (4.05) | (0.63) | (0.46) |

*, **, *** represent statistical significance at the 10%, 5%, and 1% levels, respectively.

This Table presents F-statistics testing the joint significance of the effects listed in the first column – YEAR, FIRM, IND, or LOC (MSA), and R-squared for fixed effect models. Panel A presents the results with ETR as the dependent variable, Panel B with CETR as the dependent variable, Panel C with UTB as the dependent variable, and Panel D with CETR5 as the dependent variable.

Each column represents a regression nested within the first-stage model:

$$ETR_{it}(CETR_{it}/UTB_{it}/CETR5_{it-4\sim t}) = \alpha_0 + \Sigma_l \alpha_l LOC_l + \Sigma_t \alpha_t YEAR_t + \Sigma_i \alpha_i FIRM_i + \alpha_{str} STR_{it} + \Sigma_k \alpha_k CONTROL_{it}^k + \varepsilon_{it}$$

Model 1 includes only an intercept and the vector of time-varying firm-level controls. Models 2, 3, 4 and 5 are regressions, which include only one set of fixed effects (indicator variables) – YEAR, FIRM, IND, or LOC (MSA), but include no controls for the other effects (except the effects of the time-varying firm-level controls, which are included in all models). Model 6 includes all fixed effects except for industry. All variables are defined in Panel A of Table 7.

**Table 4: Significance Levels of MSA Fixed Effects**

| | $\mathbf{ETR}_{it}$ | | $\mathbf{CETR}_{it}$ | | $\mathbf{UTB}_{it}$ | | $\mathbf{CETR5}_{it-4\sim t}$ | |
|---|---|---|---|---|---|---|---|---|
| **Level** | N | % | N | % | N | % | N | % |
| 1 percent | 23 | 16.2 | 122 | 85.9 | 31 | 29.8 | 42 | 45.7 |
| 5 percent | 47 | 33.1 | 11 | 7.8 | 19 | 18.3 | 15 | 16.3 |
| 10 percent | 21 | 14.8 | 3 | 2.1 | 15 | 14.4 | 7 | 7.6 |
| Not significant | 51 | 35.9 | 6 | 4.2 | 39 | 37.5 | 28 | 30.4 |
| Total | 142 | 100 | 142 | 100 | 104 | 100 | 92 | 100 |

This Table presents the number and percentage of individual MSAs with statistically significant MSA fixed effects from Model 6 in Table 3, Panels A, B, C and D (for ETR, CETR, UTB, and CETR5 respectively). The calculations are performed at the 1%, 5% and 10% significance level, respectively.

**Table 5: MSA Fixed Effects Variation: Length of Localisation and Geographical Diversification**

**Panel A: Length of Localisation and Location Fixed Effects**

| | $\mathbf{ETR}_{it}$ | | $\mathbf{CETR}_{it}$ | | $\mathbf{UTB}_{it}$ | | $\mathbf{CETR5}_{it-4\sim t}$ | |
|---|---|---|---|---|---|---|---|---|
| | Long | Short | Long | Short | Long | Short | Long | Short |
| Joint Significance (F-statistics) | | | | | | | | |
| $\mathrm{YEAR}_t$ | 7.61*** | 1.69** | 11.25*** | 2.37*** | 4.23*** | 0.56 | 31.59*** | 1.58** |
| $\mathrm{FIRM}_i$ | 3.26*** | 2.40*** | 2.71*** | 2.03*** | 17.28*** | 11.29*** | 11.12*** | 13.88*** |
| $\mathrm{LOC}_l$ (MSA) | 2.63*** | 1.44** | 2.68*** | 1.72*** | 5.00*** | 1.45* | 6.62*** | 3.29*** |
| Controls | YES | YES | YES | YES | YES | YES | YES | YES |
| N | 26,480 | 2,813 | 26,480 | 2,813 | 15,617 | 1,180 | 19,447 | 1,260 |
| NYEARS | 24 | 24 | 24 | 24 | 11 | 11 | 24 | 24 |
| NFIRMS | 3,858 | 1,562 | 3,858 | 1,562 | 2,823 | 665 | 2,683 | 668 |
| NLOCS | 212 | 165 | 212 | 165 | 158 | 100 | 162 | 102 |
| RSQ | 0.410 | 0.798 | 0.378 | 0.770 | 0.806 | 0.951 | 0.717 | 0.962 |

**Panel B: Geographical Diversification and Location Fixed Effects**

| | $\mathbf{ETR}_{it}$ | | $\mathbf{CETR}_{it}$ | | $\mathbf{UTB}_{it}$ | | $\mathbf{CETR5}_{it-4\sim t}$ | |
|---|---|---|---|---|---|---|---|---|
| | Yes | No | Yes | No | Yes | No | Yes | No |
| Joint Significance (F-statistics) | | | | | | | | |
| $\mathrm{YEAR}_t$ | 4.40*** | 1.24 | 8.58*** | 3.79*** | 2.57*** | 4.63*** | 30.15*** | 5.10*** |
| $\mathrm{FIRM}_i$ | 3.11*** | 2.72*** | 2.65*** | 1.81*** | 12.63*** | 11.78*** | 9.51*** | 7.45*** |
| $\mathrm{LOC}_l$ (MSA) | 2.08*** | 3.08*** | 1.87*** | 2.54*** | 3.63*** | 0.74 | 4.76*** | 2.68*** |
| Controls | YES | YES | YES | YES | YES | YES | YES | YES |
| N | 16,572 | 9,404 | 16,572 | 9,404 | 9,183 | 3,555 | 9,037 | 5,434 |
| NYEARS | 20 | 20 | 20 | 20 | 8 | 8 | 20 | 20 |
| NFIRMS | 3,402 | 3,011 | 3,402 | 3,011 | 2,104 | 1,381 | 2,301 | 1,605 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| NLOCS | 198 | 198 | 198 | 198 | 147 | 130 | 156 | 146 |
| RSQ | 0.490 | 0.621 | 0.457 | 0.529 | 0.827 | 0.888 | 0.750 | 0.822 |

*, **, *** represent statistical significance at the 10%, 5%, and 1% levels, respectively.
This Table presents for each subset of the sample F-statistics testing the joint significance of the effects listed in the first column – YEAR, FIRM, IND, or LOC (MSA); and R-squared for each regression. Panel A presents the results for whether the length of localisation affects location fixed effects. The length of localisation is deemed to be *long* if the firm has been located in the same MSA for more than five years (or three years for UTB sample). Panel B presents the results for whether the geographical diversification of the firm affects location fixed effects. The geographical diversification is deemed to be *yes* if the firm has at least one material subsidiary in another state/country in a given year.

**Table 6: MSA Fixed Effects and MSA Characteristics**

**Panel A: Descriptive Statistics**

| | **N** | **Mean** | **Std.Dev.** | **25th Pctl** | **Median** | **75th Pctl** |
|---|---|---|---|---|---|---|
| Dependent Variables | | | | | | |
| $\hat{\alpha}_l^{ETR}$ | 120 | 0.271 | 0.228 | 0.211 | 0.285 | 0.348 |
| $\hat{\alpha}_l^{CETR}$ | 120 | 0.767 | 0.265 | 0.684 | 0.807 | 0.892 |
| $\hat{\alpha}_l^{UTB}$ | 86 | 0.021 | 0.014 | 0.012 | 0.022 | 0.027 |
| $\hat{\alpha}_l^{CETR5}$ | 77 | 0.159 | 0.107 | 0.092 | 0.172 | 0.224 |
| Independent Variables | | | | | | |
| $\text{LPOP}_l$ | 120 | 12.791 | 1.205 | 11.910 | 12.685 | 13.707 |
| $\text{LWAGES}_l$ | 120 | 10.623 | 0.183 | 10.488 | 10.609 | 10.729 |
| $\text{LGDP}_l$ | 120 | 10.718 | 0.270 | 10.551 | 10.723 | 10.872 |
| $\text{EDU}_l$ | 120 | 7.510 | 0.362 | 7.264 | 7.475 | 7.784 |
| $\text{CRIMES}_l$ | 120 | 3.436 | 1.080 | 2.728 | 3.445 | 4.024 |
| $\text{RELIGION}_l$ | 120 | 0.653 | 0.082 | 0.607 | 0.650 | 0.705 |
| $\text{LAUDITOR}_l$ | 120 | 0.656 | 0.767 | 0.000 | 0.347 | 1.409 |
| $\text{LDISTANCEIRS}_l$ | 120 | 3.711 | 1.954 | 3.483 | 4.283 | 4.840 |
| $\text{LWEATHER}_l$ | 120 | 9.696 | 0.110 | 9.595 | 9.678 | 9.797 |
| $\text{LOSTFIRMPCT}_l$ | 120 | 0.067 | 0.082 | 0.014 | 0.035 | 0.082 |

**Panel B: Correlation Matrix**

| | **Variable** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** | **11** | **12** | **13** | **14** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **1** | $\hat{\alpha}_l^{ETR}$ | 1 | | | | | | | | | | | | | |
| **2** | $\hat{\alpha}_l^{CETR}$ | **0.332** | 1 | | | | | | | | | | | | |
| **3** | $\hat{\alpha}_l^{UTB}$ | **0.089** | -0.012 | 1 | | | | | | | | | | | |
| **4** | $\hat{\alpha}_l^{CETR5}$ | **0.336** | **0.338** | -0.013 | 1 | | | | | | | | | | |
| **5** | $\text{LPOP}_l$ | **0.119** | **0.135** | **0.128** | -0.017 | 1 | | | | | | | | | |
| **6** | $\text{LWAGES}_l$ | 0.0353 | **0.180** | **0.166** | 0.051 | **0.696** | 1 | | | | | | | | |
| **7** | $\text{LGDP}_l$ | **-0.071** | **0.127** | **0.208** | -0.014 | **0.523** | **0.760** | 1 | | | | | | | |
| **8** | $\text{EDU}_l$ | 0.052 | **0.214** | -0.015 | **0.123** | **0.440** | **0.692** | **0.484** | 1 | | | | | | |
| **9** | $\text{CRIMES}_l$ | **-0.084** | **-0.125** | -0.008 | **-0.103** | **-0.204** | **-0.361** | **-0.248** | **-0.267** | 1 | | | | | |
| **10** | $\text{RELIGION}_l$ | 0.017 | -0.033 | **-0.101** | **-0.127** | **-0.075** | **-0.238** | **-0.068** | **-0.243** | **0.385** | 1 | | | | |
| **11** | $\text{LAUDITOR}_l$ | -0.056 | **0.126** | **0.145** | **0.115** | **0.331** | **0.215** | **0.233** | **0.259** | **-0.172** | **-0.320** | 1 | | | |
| **12** | $\text{LDISTANCEIRS}_l$ | **-0.094** | 0.001 | -0.072 | -0.033 | **-0.528** | **-0.408** | **-0.239** | **-0.220** | **0.243** | **0.107** | **-0.194** | 1 | | |
| **13** | $\text{LWEATHER}_l$ | **0.145** | 0.050 | **-0.190** | -0.041 | **0.177** | 0.0241 | -0.050 | **-0.298** | **0.253** | **0.240** | **-0.245** | -0.057 | 1 | |
| **14** | $\text{LOSTFIRMPCT}_l$ | -0.049 | **-0.130** | **-0.134** | **0.122** | **-0.608** | **-0.477** | **-0.445** | **-0.272** | 0.0585 | 0.0116 | **-0.247** | **0.210** | 0.010 | 1 |

**Panel C: MSA Fixed Effects on MSA Characteristics**

| **Variables** | **ETR** | **CETR** | **UTB** | **CETR5** |
|---|---|---|---|---|
| | **Fixed Effects** | **Fixed Effects** | **Fixed Effects** | **Fixed Effects** |
| Intercept | −1.282 | −4.774** | 0.230 | −0.193 |
| | (−0.52) | (−2.14) | (1.53) | (−0.16) |
| $LPOP_l$ | 0.022 | −0.017 | 0.002 | −0.003 |
| | (0.71) | (−0.56) | (0.92) | (−0.16) |
| $LWAGES_l$ | −0.170 | −0.013 | 0.018 | 0.021 |
| | (−0.60) | (−0.05) | (1.01) | (0.15) |
| $LGDP_l$ | −0.145 | −0.031 | 0.007 | −0.049 |
| | (−0.90) | (−0.25) | (1.00) | (−0.69) |
| $EDU_l$ | 0.120 | 0.193*** | −0.015** | 0.053 |
| | (1.06) | (2.70) | (−2.18) | (1.10) |
| $CRIMES_l$ | −0.033 | −0.037 | 0.002 | −0.005 |
| | (−1.24) | (−1.50) | (1.09) | (−0.31) |
| $RELIGION_l$ | 0.069 | 0.205 | −0.029 | −0.092 |
| | (0.28) | (0.52) | (−1.36) | (−0.47) |
| $LAUDITOR_l$ | −0.021 | 0.044 | 0.001 | 0.020 |
| | (−0.51) | (1.43) | (0.57) | (1.18) |
| $LDISTANCEIRS_l$ | −0.005 | 0.013 | 0.000 | −0.001 |
| | (−0.45) | (0.81) | (0.06) | (−0.09) |
| $LWEATHER_l$ | 0.397** | 0.486** | −0.038** | 0.036 |
| | (2.03) | (2.57) | (−2.22) | (0.31) |
| $LOSTFIRMPCT_l$ | −0.198 | −0.341 | 0.012 | 0.403 |
| | (−0.42) | (−0.62) | (0.37) | (1.58) |
| N | 120 | 120 | 86 | 77 |

| | | | | |
|---|---|---|---|---|
| RSQ | 0.084 | 0.098 | 0.168 | 0.076 |
| F-test | 1.17 | 1.65 | 2.43** | 1.00 |

*, **, *** represent statistical significance at the 10%, 5%, and 1% levels, respectively.

Panel A reports descriptive statistics for regression of MSA fixed effects and MSA characteristics. Panel B reports Pearson correlation coefficients for the variables used in Equation (2). All variables are defined in Panel B of Table 7. **Bolded** coefficients denote significance at the 5% level or less using a two-sided test. Panel C reports the results from testing Equation (2) with robust standard errors:

$$\hat{\alpha}_l^{ETR}(\hat{\alpha}_l^{CETR}/\hat{\alpha}_l^{UTB}/\hat{\alpha}_l^{CETR5}) = \beta_0 + \beta_1 LPOP_l + \beta_2 LWAGES_l + \beta_3 LGDP_l + \beta_4 EDUC_l + \beta_5 CRIMES_l + \beta_6 RELIGION_l + \beta_7 LAUDITOR_l + \beta_8 LDISTANCEIRS_l + \beta_9 LWEATHER_l + \beta_{10} LOSTFIRMPCT_l + \epsilon_l$$

The dependent variables are coefficients estimated for MSA $l$'s fixed effect when the dependent variables in Equation (1) are ETR, CETR, UTB, and CETR5 respectively. The independent variables include a number of MSA characteristics.

**Table 7: Variable Definitions**

**Panel A: Variable Definitions for Equation (1)**

| Variable | Definition |
|---|---|
| **Dependent variables:** | |
| $\mathrm{ETR}_{it}$ | The financial accounting (i.e. GAAP) effective tax rate, defined as total income tax expense (TXT) divided by pre-tax book income (PI) before special items (SPI) |
| $\mathrm{CETR}_{it}$ | The cash effective tax rate, defined as cash tax paid (TXPD) divided by pre-tax book income (PI) before special items (SPI) |
| $\mathrm{UTB}_{it}$ | The FIN 48 tax reserve, measured as the end of year unrecognised tax benefits (TXTUBEND) scaled by total assets (AT) |
| $\mathrm{CETR5}_{it-4\sim t}$ | The long-run average cash effective tax rate, defined as the sum of cash taxes paid (TXPD) in years t, t-1, t-2, t-3, and t-4, divided by pre-tax book income (PI) before special items (SPI) in years t, t-1, t-2, t-3, and t-4 |
| **Fixed effects:** | |
| $\mathrm{LOC}_l$ (MSA fixed effects) | A set of indicator variables indicating the Metropolitan Statistical Area (MSA) of the firm's headquarters |
| $\mathrm{FIRM}_i$ (firm fixed effects) | A set of indicator variables for each firm (GVKEY) |
| $\mathrm{YEAR}_t$ (year fixed effects) | A set of indicator variables indicating the financial year (FYEAR) of the observation |
| $\mathrm{IND}_i$ (industry fixed effects) | A set of indicator variables indicating the industry membership (four-digit SIC code) of the firm |
| **Control variables:** | |
| $\mathrm{SIZE}_{it}$ | Firm size, measured as the natural log of total assets (AT) |
| $\mathrm{ROA}_{it}$ | Return on assets, measured as pre-tax book income (PI) divided by lagged total assets (AT) |
| $\mathrm{NOL}_{it}$ | A dummy variable coded as one if the firm reports a positive tax loss carry forward (TLCF) and zero otherwise |
| $\Delta\mathrm{NOL}_{it}$ | Change in tax loss carry forward (ΔTLCF), scaled by lagged total assets (AT); when missing, reset to 0 |
| $\mathrm{LEV}_{it}$ | Leverage, measured as the sum of long-term debt (DLTT) and long-term debt in current liabilities (DLC) divided by total assets (AT) |

| | |
|---|---|
| $FI_{it}$ | Pre-tax foreign income (PIFO), scaled by lagged total assets (AT); when missing, reset to 0 |
| $GPPE_{it}$ | Gross property, plant, and equipment (PPEGT), scaled by total assets (AT) |
| $INTANG_{it}$ | Intangible assets (INTAN), scaled by total assets (AT) |
| $EQINC_{it}$ | Equity income in earnings (ESUB), scaled by lagged assets (AT) |
| $MTB_{it}$ | Market-to-book ratio, measured as market value of equity (PRCC_F*CSHO), scaled by book value of equity (CEQ) |
| $R\&D_{it}$ | Research and development expense (XRD), scaled by net sales (SALE); when missing, reset to 0 |
| $CASH_{it}$ | Cash and cash equivalents (CHE), scaled by total assets (AT) |
| $CAPEX_{it}$ | Capital expenditures (CAPX), scaled by gross property, plant, and equipment (PPEGT) |
| $ADV_{it}$ | Advertising expense (XAD), scaled by net sales (SALE); when missing, reset to 0 |
| $SG\&A_{it}$ | Selling, general, and administrative expense (XSGA), scaled by net sales (SALE); when missing, reset to 0 |
| $\Delta SALES_{it}$ | The annual percentage change in net sales (($SALE_t$ / $SALE_{t-1}$) – 1) |
| $INSTPERC_{it}$ | The percentage of nontransient institutional investors, following Bushee (1998); when missing, reset to 0 |
| $MKTPRES_{it}$ | Count of the number of consecutive nonnegative changes in split-adjusted quarterly earnings per share relative to the same quarter from the prior year, following Myers, Myers and Skinner (2007) |
| $ANALYST_{it}$ | The natural logarithm of the number of analyst estimates reported before the end of the fiscal year; when missing, reset to 0 |
| $ALOCAL_{it}$ | The average number of firms in the fiscal year followed by each analyst in the same MSA; when missing, reset to 0 |
| $NUMDIRS_{it}$ | The natural logarithm of the number of total directors sitting on the Board |
| $PCTINDEP_{it}$ | The percentage of independent directors to total directors sitting on the Board |
| $GENDER_{it}$ | The proportion of male directors sitting on the Board |
| $CFOBOD_{it}$ | An indicator variable equals 1 if the CFO is on the Board of Directors and 0 otherwise |
| $CEOCHAIR_{it}$ | An indicator variable that equals 1 if the CEO is also Chair of the Board and 0 otherwise |
| $CEOCOMP_{it}$ | The natural logarithm of CEO total compensation |
| $STR_{it}$ | The highest marginal rate as reported in the state corporate income tax schedule in a given year |

**Panel B: Variable Definitions for Equation (2)**

| Variable | Definition |
|---|---|
| **Dependent variables:** | |
| $\hat{\alpha}_l^{ETR}$ | The coefficients estimated when ETR is the dependent variable in Equation (1) |
| $\hat{\alpha}_l^{CETR}$ | The coefficients estimated when CETR is the dependent variable in Equation (1) |

| | |
|---|---|
| $\hat{\alpha}_l^{UTB}$ | The coefficients estimated when UTB is the dependent variable in Equation (1) |
| $\hat{\alpha}_l^{CETR5}$ | The coefficients estimated when CETR5 is the dependent variable in Equation (1) |
| **Independent variables:** | |
| $\text{LPOP}_l$ | Natural logarithm of the estimated size of the average workforce population for the MSA |
| $\text{LWAGES}_l$ | Natural logarithm of the weighted-average personal income for the employed workforce in the MSA, calculated using personal income (INCWAGE) weighted by sample weights (PERWT) from the IPUMS |
| $\text{LGDP}_l$ | Natural logarithm of the average GDP per capita in the MSA, calculated using the average total GDP index divided by the average total population in the MSA |
| $\text{EDU}_l$ | The weighted-average education level of the workforce population in the MSA, calculated using education levels (EDUC) weighted by sample weights from the IPUMS |
| $\text{CRIMES}_l$ | The average total crime rate in the MSA, calculated by dividing the number of reported crimes by the total population and multiplying the result by 100 |
| $\text{RELIGION}_l$ | The proportion of Gallup poll respondents in each state, matched to each MSA, who indicate that religion is an important part of their daily life |
| $\text{LAUDITOR}_l$ | The presence of BIG4 audit firms' local offices in the MSA, measured by the average of the natural logarithm of one plus the number of all BIG4 firms' local offices |
| $\text{LDISTANCEIRS}_l$ | Natural logarithm of the geodesic distance, reported in miles, between the firm's headquarters and the closest IRS territory manager's office |
| $\text{LWEATHER}_l$ | Natural logarithm of the state-level average daily sunlight (insolation or solar radiation) matched to each MSA, reported in kilojoules per square meter (KJ/m2) |
| $\text{LOSTFIRMPCT}_l$ | The percentage of firms that report a negative pre-tax book income out of all firms in each MSA |

**Fig. 1: Spatial Distribution of MSA Fixed Effect Coefficients**

**Fig. 1.1 – spatial distribution of the inverted location fixed effect coefficients estimated in Table 3, Panel A (for ETR). A darker shade indicates greater corporate tax avoidance.**

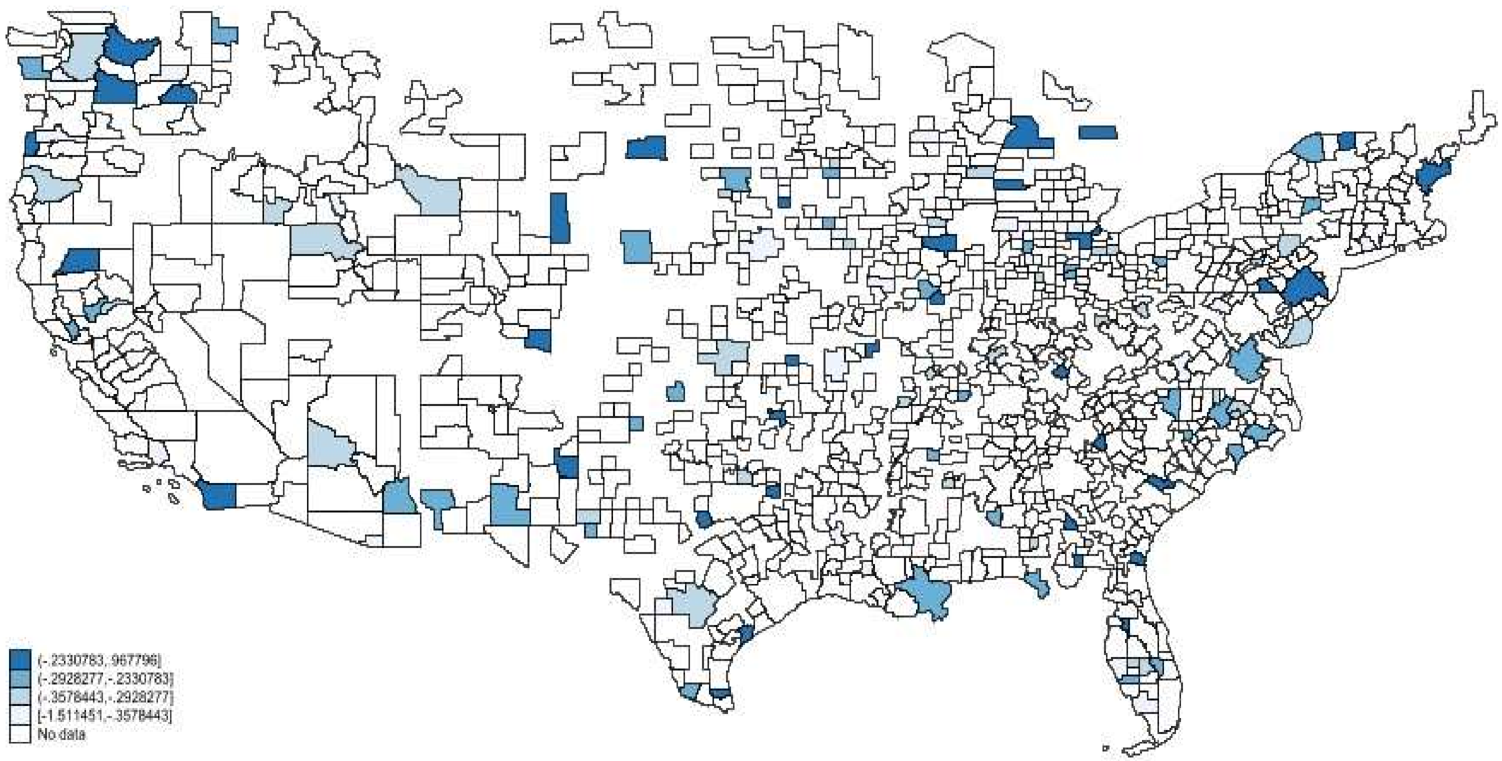


**Fig. 1.2 – spatial distribution of the inverted location fixed effect coefficients estimated in Table 3, Panel B (for CETR). A darker shade indicates greater corporate tax avoidance.**

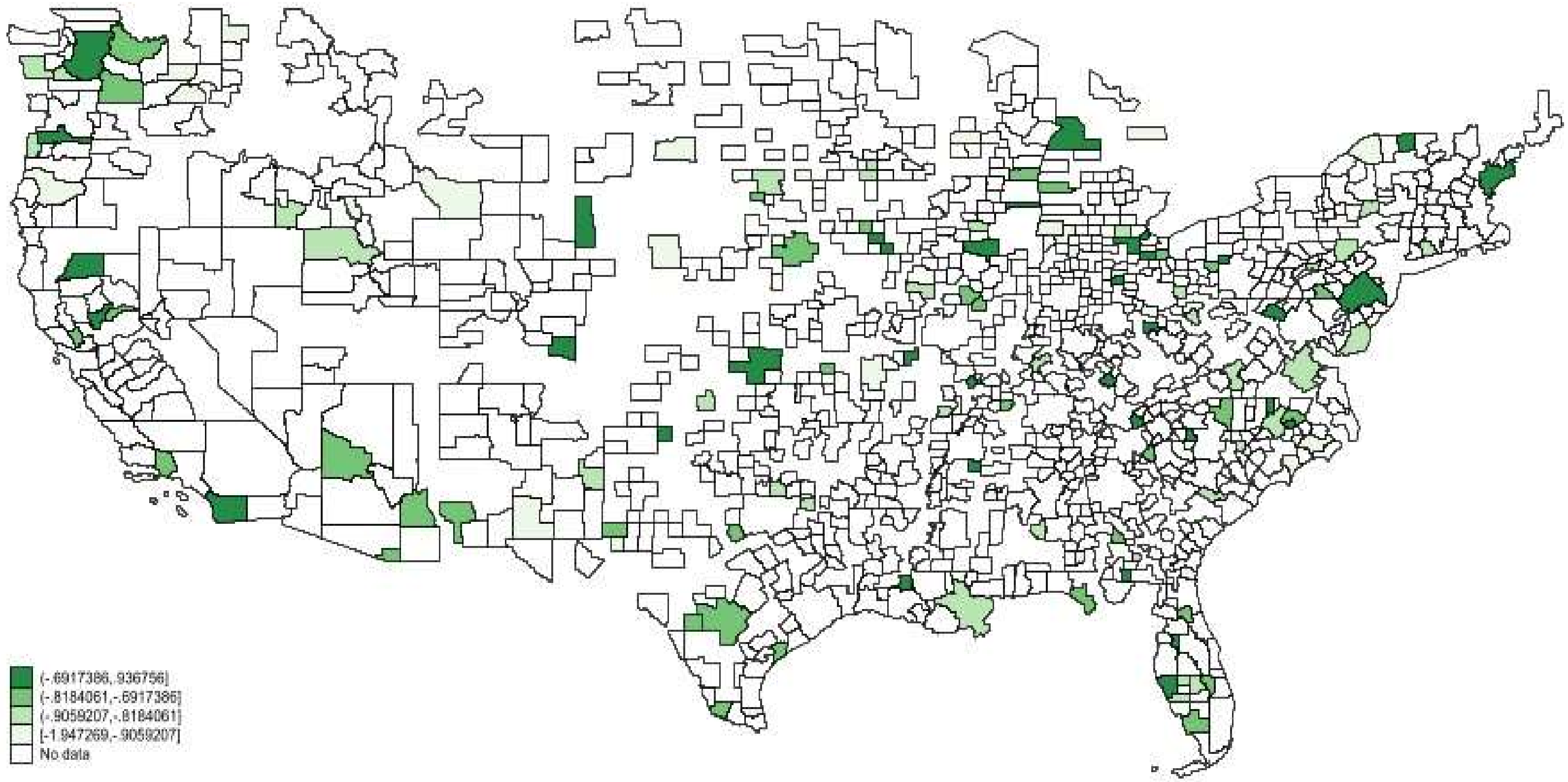


**Fig. 1.3 – spatial distribution of the location fixed effect coefficients estimated in Table 3, Panel C (for UTB). A darker shade indicates more uncertain tax positions.**

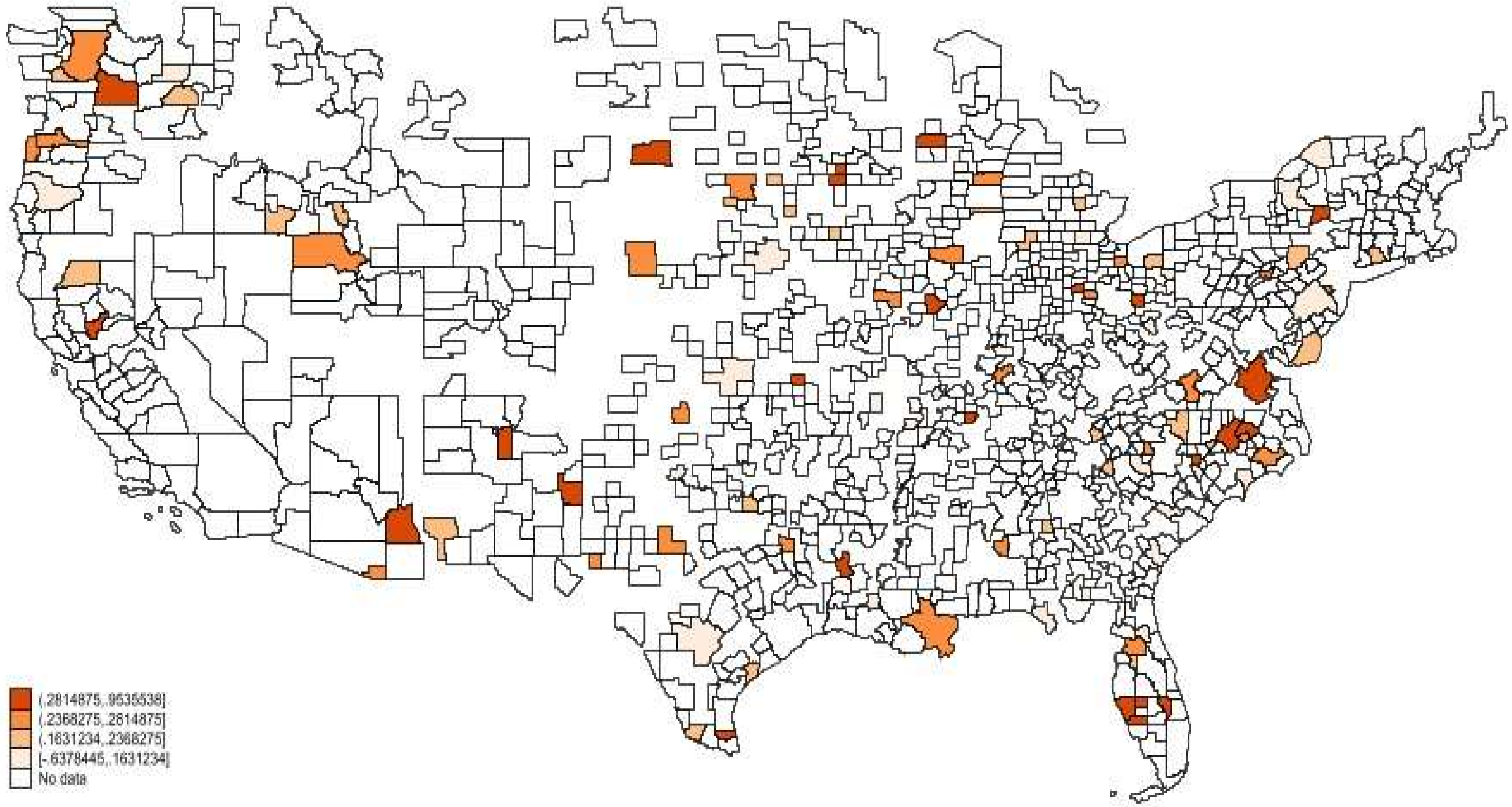
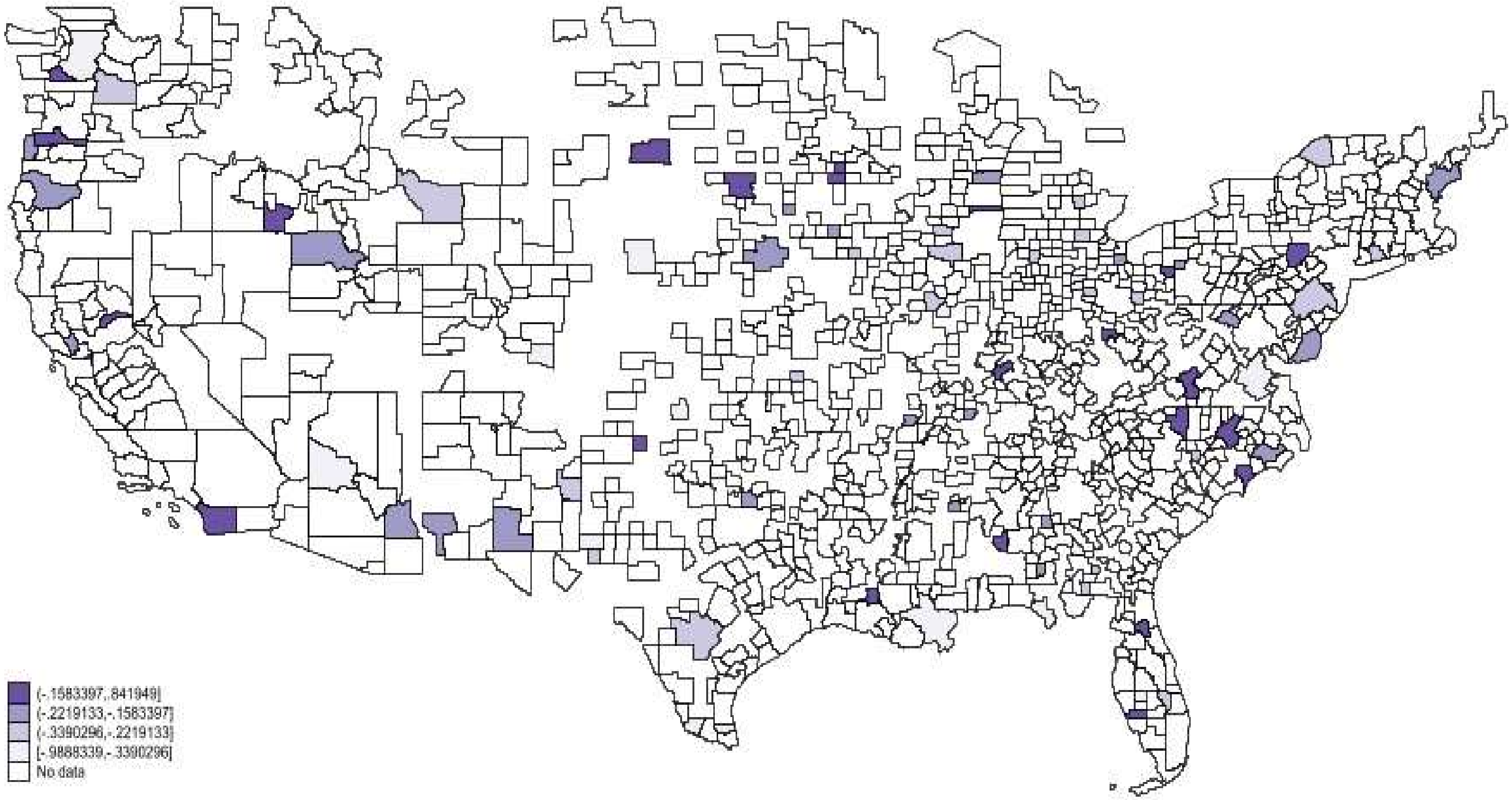


**Fig. 1.4 – spatial distribution of the inverted location fixed effect coefficients estimated in Table 3, Panel D (for CETR5). A darker shade indicates greater long-run tax avoidance.**



## Supplementary Appendix: Robustness Tests

### 1. Including loss observations

In our main sample, we exclude firm-year observations with negative pre-tax income following prior studies (Dyreng et al., 2010; Kubick et al., 2017), which results in a significant reduction in the sample size. This section tests the sensitivity of our results to including those loss observations. Specifically, we employ a measure of corporate tax avoidance that uses the market value of assets in place of pre-tax income in its denominator. Our measure, ΔCETR, is calculated as the difference between cash taxes paid and the product of pre-tax income and the corporate statutory tax rate, scaled by market value of assets (Henry and Sansing, 2018). ΔETR uses GAAP tax expense instead of cash taxes paid. ΔCETR5 is ΔCETR measured in the long-run with both the numerator and the denominator summed over a five-year period. Including loss firms increases the sample size to between 42,796 to 55,226 firm-year observations. The results presented in Table SA1 show that MSA fixed effects are jointly significant in explaining ΔETR and ΔCETR and ΔCETR5 (F-statistic = 2.81, 2.47, and 8.25, respectively).[9] These effects are still significant when other fixed effects are controlled in the model (F-statistic = 2.26, 2.16 and 3.37, respectively). This is consistent with our main findings and it suggests that locations fixed effects are present in both profitable and loss firms.

**Table SA1: Location Fixed Effects on Tax Avoidance: Including Loss Observations**

**Panel A: MSA Fixed Effects on ΔETR ($\Delta \text{ETR}_{it}$)**

| | Model 1 | Model 2 | Model 3 | Model 4 | Model 5 | Model 6 |
|---|---|---|---|---|---|---|
| Joint Significance (F-statistics) | | | | | | |
| $\text{YEAR}_t$ | | 50.07*** | | | | 43.03*** |
| $\text{FIRM}_i$ | | | 3.24*** | | | 3.23*** |
| $\text{IND}_i$ | | | | 2.89*** | | ---------- |
| $\text{LOC}_l$ (MSA) | | | | | 2.47*** | 2.16*** |

[9] Note that we did not exclude loss firms for the UTB sample in our main analysis. Therefore, the location fixed effects on the UTB of both profitable and loss firms have already been estimated and the results are presented in Table 3 (see the main results).

| | | | | | | |
|---|---|---|---|---|---|---|
| N | 55,226 | 55,226 | 55,226 | 55,226 | 55,226 | 55,226 |
| NYEARS | | 24 | | | | 24 |
| NFIRMS | | | 7,559 | | | 7,559 |
| NINDS | | | | 371 | | ---------- |
| NLOCS | | | | | 330 | 330 |
| RSQ | 0.274 | 0.289 | 0.521 | 0.288 | 0.285 | 0.535 |

**Panel B: MSA Fixed Effects on ΔCETR ($\Delta\mathbf{CETR}_{it}$)**

| | **Model 1** | **Model 2** | **Model 3** | **Model 4** | **Model 5** | **Model 6** |
|---|---|---|---|---|---|---|
| Joint Significance (F-statistics) | | | | | | |
| $\text{YEAR}_t$ | | 56.02*** | | | | 51.81*** |
| $\text{FIRM}_i$ | | | 3.55*** | | | 3.54*** |
| $\text{IND}_i$ | | | | 3.24*** | | ---------- |
| $\text{LOC}_l$ (MSA) | | | | | 2.81*** | 2.26*** |
| N | 55,226 | 55,226 | 55,226 | 55,226 | 55,226 | 55,226 |
| NYEARS | | 24 | | | | 24 |
| NFIRMS | | | 7,559 | | | 7,559 |
| NINDS | | | | 371 | | ---------- |
| NLOCS | | | | | 330 | 330 |
| RSQ | 0.303 | 0.318 | 0.554 | 0.318 | 0.314 | 0.569 |

**Panel C: MSA Fixed Effects on ΔCETR5 ($\Delta\mathbf{CETR5}_{it-4\sim t}$)**

| | **Model 1** | **Model 2** | **Model 3** | **Model 4** | **Model 5** | **Model 6** |
|---|---|---|---|---|---|---|
| Joint Significance (F-statistics) | | | | | | |
| $\text{YEAR}_t$ | | 37.99*** | | | | 95.73*** |
| $\text{FIRM}_i$ | | | 10.87*** | | | 10.92*** |
| $\text{IND}_i$ | | | | 16.61*** | | ---------- |
| $\text{LOC}_l$ (MSA) | | | | | 8.25*** | 3.37*** |

| | | | | | | |
|---|---|---|---|---|---|---|
| N | 42,796 | 42,796 | 42,796 | 42,796 | 42,796 | 42,796 |
| NYEARS | | 24 | | | | 24 |
| NFIRMS | | | 5,606 | | | 5,606 |
| NINDS | | | | 365 | | ---------- |
| NLOCS | | | | | 314 | 314 |
| RSQ | 0.211 | 0.226 | 0.701 | 0.310 | 0.257 | 0.722 |

*, **, *** represent statistical significance at the 10%, 5%, and 1% levels, respectively.

This Table presents F-statistics testing the joint significance of the effects listed in the first column – YEAR, FIRM, IND, or LOC (MSA), and R-squared for fixed effect models. Panel A presents the results with the modified effective tax rate measure ΔETR as the dependent variable, Panel B with the Henry and Sansing (2018) measure ΔCETR as the dependent variable, and Panel C with the modified ΔCETR5 as the dependent variable. ΔCETR is the difference between cash taxes paid, adjusted for the change in tax refunds receivable, and the product of pretax income and the statutory tax rate, scaled by market value of assets, i.e. (TXPD-($TXR_t$-$TXR_{t-1}$)-0.35*PI)/(AT+PRCC_Q*CSHO-SEQ). ΔETR uses GAAP tax expense instead of cash taxes paid, i.e. (TXT-0.35*PI)/(AT+PRCC_Q*CSHO-SEQ). ΔCETR5 is ΔCETR with both the numerator and the denominator measured over a five-year period (t-4 to t).

Each column represents a regression nested within the first-stage model:

$$\Delta ETR_{it}(\Delta CETR_{it}\ /\ \Delta CETR5_{it}) = \alpha_0 + \Sigma_l\alpha_l LOC_l + \Sigma_t\alpha_t YEAR_t + \Sigma_i\alpha_i FIRM_i + \alpha_{str}STR_{it} + \Sigma_k\alpha_k CONTROL_{it}^k + \varepsilon_{it}$$

Model 1 includes only an intercept and the vector of time-varying firm-level controls. Models 2, 3, 4 and 5 are regressions, which include only one set of fixed effects (indicator variables) – YEAR, FIRM, IND, or LOC (MSA), but include no controls for the other effects (except the effects of the time-varying firm-level controls, which are included in all models). Model 6 includes all fixed effects except for industry.

## 2. Alternative measures for geographic units

Our main findings of the location fixed effect on corporate tax avoidance are based on the basic geographic unit, MSAs. We also consider alternative measures of geographic unit, including states, counties, and the zip codes of the firms' headquarters. These three additional geographic units tested here reflect different ways of grouping geographical areas for administrative and political reasons. As reported in Table SA2, models 5, 6 and 7 replace the MSAs in Equation (1) with the new measures for geographical unit, i.e. states, counties, and the zip codes, respectively. Similarly, models 8, 9 and 10 include year, firm and location fixed effects with those new geographical unit proxies substituting MSAs. We consistently find that firm locations have jointly significant effects on corporate tax avoidance, regardless of the specific unit of measurement we use in the analysis. The F-statistics are usually the largest for state fixed effects and smaller for county and zip code fixed effects, both in the full models and when each location fixed effect is tested individually. This is because both county and zip code are smaller geographic units that contain the least number of firms and therefore the location fixed effects overlap more with the firm fixed effects. The R-squared is the largest for the zip code fixed effects, which is unsurprising given that R-squared increases as more predictors are included. Overall, the results from estimating Equation (1) are robust to using different geographic units as the main source of variation.

**Table SA2: Location Fixed Effects on Tax Avoidance: Alternative Geographical Units**

**Panel A: Geographical Fixed Effects on Effective Tax Rates ($ETR_{it}$)**

| | Model 1 | Model 2 | Model 3 | Model 4 | Model 5 | Model 6 | Model 7 | Model 8 | Model 9 | Model 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| $YEAR_t$ | | 7.23*** | | | | | | 7.67*** | 7.51*** | 6.86*** |
| $FIRM_i$ | | | 3.01*** | | | | | 3.02*** | 3.08*** | 3.13*** |
| $IND_i$ | | | | 3.20*** | | | | ---------- | ---------- | ---------- |
| $LOC_l$ (state) | | | | | 2.29*** | | | 2.26*** | | |
| $LOC_l$ (county) | | | | | | 1.68*** | | | 2.23*** | |
| $LOC_l$ (zipcode) | | | | | | | 2.20*** | | | 2.21*** |
| CONTROLS | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES |
| N | 29,293 | 29,293 | 29,293 | 29,293 | 29,293 | 29,293 | 29,293 | 29,293 | 29,293 | 29,293 |
| NYEARS | | 24 | | | | | | 24 | 24 | 24 |
| NFIRMS | | | 5,197 | | | | | 5,197 | 5,197 | 5,197 |
| NINDS | | | | 367 | | | | ---------- | ---------- | ---------- |
| NLOCS | | | | | 50 | 557 | 2,769 | 50 | 557 | 2,769 |
| RSQ | 0.071 | 0.076 | 0.437 | 0.107 | 0.074 | 0.093 | 0.244 | 0.443 | 0.453 | 0.502 |

**Panel B: Geographical Fixed Effects on Cash Effective Tax Rates ($CETR_{it}$)**

| | Model 1 | Model 2 | Model 3 | Model 4 | Model 5 | Model 6 | Model 7 | Model 8 | Model 9 | Model 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| $YEAR_t$ | | 11.59*** | | | | | | 12.21*** | 12.06*** | 10.76*** |
| $FIRM_i$ | | | 2.54*** | | | | | 2.54*** | 2.51*** | 2.52*** |
| $IND_i$ | | | | 2.98*** | | | | ---------- | ---------- | ---------- |
| $LOC_l$ (state) | | | | | 2.67*** | | | 1.55** | | |
| $LOC_l$ (county) | | | | | | 2.17*** | | | 1.61*** | |
| $LOC_l$ (zipcode) | | | | | | | 2.05*** | | | 1.84*** |
| CONTROLS | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES |
| N | 29,293 | 29,293 | 29,293 | 29,293 | 29,293 | 29,293 | 29,293 | 29,293 | 29,293 | 29,293 |
| NYEARS | | 24 | | | | | | 24 | 24 | 24 |
| NFIRMS | | | 5,197 | | | | | 5,197 | 5,197 | 5,197 |
| NINDS | | | | 367 | | | | ---------- | ---------- | ---------- |
| NLOCS | | | | | 50 | 557 | 2,769 | 50 | 557 | 2,769 |
| RSQ | 0.075 | 0.083 | 0.403 | 0.109 | 0.079 | 0.104 | 0.238 | 0.411 | 0.419 | 0.464 |

**Panel C: Geographical Fixed Effects on FIN 48 Tax Reserve ($UTB_{it}$)**

| | Model 1 | Model 2 | Model 3 | Model 4 | Model 5 | Model 6 | Model 7 | Model 8 | Model 9 | Model 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| $YEAR_t$ | | 10.65*** | | | | | | 4.25*** | 4.39*** | 4.34*** |
| $FIRM_i$ | | | 16.70*** | | | | | 15.85*** | 15.26*** | 17.34*** |
| $IND_i$ | | | | 5.85*** | | | | ---------- | ---------- | ---------- |
| $LOC_l$ (state) | | | | | 20.65*** | | | 4.23*** | | |
| $LOC_l$ (county) | | | | | | 7.60*** | | | 2.73*** | |
| $LOC_l$ (zipcode) | | | | | | | 4.57*** | | | 3.62*** |
| CONTROLS | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES |
| N | 18,925 | 18,925 | 18,925 | 18,925 | 18,925 | 18,925 | 18,925 | 18,925 | 18,925 | 18,925 |
| NYEARS | | 11 | | | | | | 11 | 11 | 11 |
| NFIRMS | | | 3,322 | | | | | 3,322 | 3,322 | 3,322 |
| NINDS | | | | 326 | | | | ---------- | ---------- | ---------- |
| NLOCS | | | | | 44 | 375 | 1,814 | 44 | 375 | 1,814 |
| RSQ | 0.094 | 0.099 | 0.801 | 0.178 | 0.134 | 0.197 | 0.390 | 0.804 | 0.807 | 0.831 |

**Panel D: Geographical Fixed Effects on 5-year Cash Effective Tax Rates ($CETR5_{it-4\sim t}$)**

| | Model 1 | Model 2 | Model 3 | Model 4 | Model 5 | Model 6 | Model 7 | Model 8 | Model 9 | Model 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| $YEAR_t$ | | 26.56*** | | | | | | 32.34*** | 32.11*** | 30.32*** |
| $FIRM_i$ | | | 10.01*** | | | | | 10.10*** | 10.11*** | 9.45*** |
| $IND_i$ | | | | 9.46*** | | | | ---------- | ---------- | ---------- |
| $LOC_l$ (state) | | | | | 7.93*** | | | 5.20*** | | |
| $LOC_l$ (county) | | | | | | 4.59*** | | | 4.20*** | |
| $LOC_l$ (zipcode) | | | | | | | 5.36*** | | | 3.58*** |
| CONTROLS | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES |
| N | 20,707 | 20,707 | 20,707 | 20,707 | 20,707 | 20,707 | 20,707 | 20,707 | 20,707 | 20,707 |
| NYEARS | | 24 | | | | | | 24 | 24 | 24 |
| NFIRMS | | | 3,248 | | | | | 3,248 | 3,248 | 3,248 |
| NINDS | | | | 355 | | | | ---------- | ---------- | ---------- |
| NLOCS | | | | | 47 | 419 | 2,046 | 47 | 419 | 2,046 |
| RSQ | 0.169 | 0.192 | 0.710 | 0.287 | 0.184 | 0.228 | 0.477 | 0.725 | 0.732 | 0.761 |

*, **, *** represent statistical significance at the 10%, 5%, and 1% levels, respectively.

This Table presents F-statistics testing the joint significance of the effects listed in the first column – YEAR, FIRM, or IND, and alternative measures of LOC (state, county or zip code); and R-squared for fixed effect models. The estimated coefficients for control variables are omitted for brevity. Panel A presents the results with ETR as the dependent variable, Panel B with CETR as the dependent variable, Panel C with UTB as the dependent variable, and Panel D with CETR5 as the dependent variable. Each column represents a regression nested within the first-stage model:

$$ETR_{it}(CETR_{it}/UTB_{it}/CETR5_{it-4\sim t}) = \alpha_0 + \Sigma_l \alpha_l LOC_l + \Sigma_t \alpha_t YEAR_t + \Sigma_i \alpha_i FIRM_i + \alpha_{str} STR_{it} + \Sigma_k \alpha_k CONTROL_{it}^k + \varepsilon_{it}$$

Model 1 includes only the vector of time-varying firm-level controls. Models 2, 3, 4 and 5 are regressions, which include only one set of effects (indicator variables) – YEAR, FIRM, IND, or LOC (state, county or zip code), but include no controls for the other effects (except the effects of the time-varying firm-level controls, which are included in all models). The final three models, Models 8, 9 and 10, are tests of each of the geographical effects in the presence of the firm and year effects.

## 3. Selection bias test: location decision

In our main specification, we control for firm fixed effects. That leads to a limitation that our results rely on firms that have changed MSA locations. There is a concern that firms' change of location can be endogenously decided and correlated with geographic factors and tax-related incentives. To address this concern, we split the sample into (1) firms that have changed headquarter locations from one MSA to another at least once during the sample period; and (2) firms that have not changed headquarter locations during the sample period. In sample one (two), there are 5,008 (24,285) observations in 164 (212) MSAs for the ETR and CETR regressions, 2,447 (16,478) observations in 114 (156) MSAs for UTB, and 3,154 (17,553) observations in 118 (162) MSAs for CETR5. Table SA3 reports results using the location-change sample. We find that MSA fixed effects are significant when included on their own across all regressions, and collectively with firm and year fixed effects, which is consistent with our main results. The F-statistics for the MSA fixed effects are generally smaller than those reported in Table 3 (see the main results), which is understandable given that the reduced sample has much fewer degrees of freedom and fewer numbers of MSAs. In Table SA4, we present results using the second sample of no-change firms. In this sample, we cannot control for firm fixed effects because they will subsume location fixed effects. In the full specification, we control for industry fixed effects instead. We continue to find location fixed effects to be jointly significant in explaining variations in tax avoidance measures. In summary, our findings suggest that location fixed effects are less likely to be affected by firms' decision on changing or not changing their headquarter locations.

**Table SA3: MSA Fixed Effects for Firms that Changed Location**

**Panel A: MSA Fixed Effects on Effective Tax Rates ($ETR_{it}$)**

| | Model 1 | Model 2 | Model 3 | Model 4 | Model 5 | Model 6 |
|---|---|---|---|---|---|---|
| Joint Significance (F-statistics) | | | | | | |
| $YEAR_t$ | | 1.12 | | | | 1.11 |
| $FIRM_i$ | | | 2.46*** | | | 2.51*** |
| $IND_i$ | | | | 2.27*** | | ---------- |
| $LOC_l$ (MSA) | | | | | 1.57*** | 1.83*** |
| Controls | YES | YES | YES | YES | YES | YES |
| N | 5,008 | 5,008 | 5,008 | 5,008 | 5,008 | 5,008 |
| NYEARS | | 24 | | | | 24 |
| NFIRMS | | | 944 | | | 944 |

| | | | | | | |
|---|---|---|---|---|---|---|
| NINDS | | | | 254 | | ---------- |
| NLOCS | | | | | 164 | 164 |
| RSQ | 0.072 | 0.076 | 0.410 | 0.172 | 0.118 | 0.450 |

**Panel B: MSA Fixed Effects on Cash Effective Tax Rates (CETR$_{it}$)**

| | **Model 1** | **Model 2** | **Model 3** | **Model 4** | **Model 5** | **Model 6** |
|---|---|---|---|---|---|---|
| Joint Significance (F-statistics) | | | | | | |
| YEAR$_t$ | | 1.61** | | | | 2.01*** |
| FIRM$_i$ | | | 2.12*** | | | 2.11*** |
| IND$_i$ | | | | 2.26*** | | ---------- |
| LOC$_l$ (MSA) | | | | | 1.61*** | 1.54*** |
| Controls | YES | YES | YES | YES | YES | YES |
| N | 5,008 | 5,008 | 5,008 | 5,008 | 5,008 | 5,008 |
| NYEARS | | 24 | | | | 24 |
| NFIRMS | | | 944 | | | 944 |
| NINDS | | | | 254 | | ---------- |
| NLOCS | | | | | 164 | 164 |
| RSQ | 0.081 | 0.087 | 0.385 | 0.180 | 0.128 | 0.424 |

**Panel C: MSA Fixed Effects on FIN 48 Tax Reserve (UTB$_{it}$)**

| | **Model 1** | **Model 2** | **Model 3** | **Model 4** | **Model 5** | **Model 6** |
|---|---|---|---|---|---|---|
| Joint Significance (F-statistics) | | | | | | |
| YEAR$_t$ | | 3.11*** | | | | 1.09 |
| FIRM$_i$ | | | 12.10*** | | | 10.81*** |
| IND$_i$ | | | | 4.55*** | | ---------- |
| LOC$_l$ (MSA) | | | | | 3.93*** | 2.41*** |
| Controls | YES | YES | YES | YES | YES | YES |
| N | 2,447 | 2,447 | 2,447 | 2,447 | 2,447 | 2,447 |
| NYEARS | | 11 | | | | 11 |
| NFIRMS | | | 485 | | | 485 |

| | | | | | | |
|---|---|---|---|---|---|---|
| NINDS | | | | 168 | | ---------- |
| NLOCS | | | | | 114 | 114 |
| RSQ | 0.093 | 0.104 | 0.774 | 0.321 | 0.239 | 0.802 |

**Panel D: MSA Fixed Effects on 5-Year Cash ETR ($CETR5_{it-4\sim t}$)**

| | Model 1 | Model 2 | Model 3 | Model 4 | Model 5 | Model 6 |
|---|---|---|---|---|---|---|
| Joint Significance (F-statistics) | | | | | | |
| $YEAR_t$ | | 3.22*** | | | | 3.48*** |
| $FIRM_i$ | | | 9.13*** | | | 9.59*** |
| $IND_i$ | | | | 7.41*** | | ---------- |
| $LOC_l$ (MSA) | | | | | 4.42*** | 5.31*** |
| Controls | YES | YES | YES | YES | YES | YES |
| N | 3,154 | 3,154 | 3,154 | 3,154 | 3,154 | 3,154 |
| NYEARS | | 24 | | | | 24 |
| NFIRMS | | | 507 | | | 507 |
| NINDS | | | | 215 | | ---------- |
| NLOCS | | | | | 118 | 118 |
| RSQ | 0.200 | 0.217 | 0.710 | 0.481 | 0.317 | 0.763 |

*, **, *** represent statistical significance at the 10%, 5%, and 1% levels, respectively.

This Table presents F-statistics testing the joint significance of the effects listed in the first column – YEAR, FIRM, IND, or LOC (MSA); and R-squared for fixed effect models. The sample only includes firms that have changed location during the sample period. Panel A presents the results with ETR as the dependent variable, Panel B with CETR as the dependent variable, Panel C with UTB as the dependent variable, and Panel D with CETR5 as the dependent variable. Each column represents a regression nested within the first-stage model:

$$ETR_{it}(CETR_{it}/UTB_{it}/CETR5_{it-4\sim t}) = \alpha_0 + \Sigma_l\alpha_l LOC_l + \Sigma_t\alpha_t YEAR_t + \Sigma_i\alpha_i FIRM_i + \alpha_{str}STR_{it} + \Sigma_k\alpha_k CONTROL_{it}^k + \varepsilon_{it}$$

Model 1 includes only an intercept and the vector of time-varying firm-level controls. Models 2, 3, 4 and 5 are regressions, which include only one set of fixed effects (indicator variables) – YEAR, FIRM, IND, or LOC (MSA), but include no controls for the other effects (except the effects of the time-varying firm-level controls, which are included in all models). Model 6 includes all fixed effects except for industry.

**Table SA4: MSA Fixed Effects for Firms that Have Not Changed Location**

**Panel A: MSA Fixed Effects on Effective Tax Rates ($ETR_{it}$)**

| | **Model 1** | **Model 2** | **Model 3** | **Model 4** | **Model 5** | **Model 6** |
|---|---|---|---|---|---|---|
| Joint Significance (F-statistics) | | | | | | |
| $YEAR_t$ | | 6.94*** | | | | 6.48*** |
| $FIRM_i$ | | | 3.17*** | | | ---------- |
| $IND_i$ | | | | 3.04*** | | 2.96*** |
| $LOC_l$ (MSA) | | | | | 1.86*** | 1.71*** |
| Controls | YES | YES | YES | YES | YES | YES |
| N | 24,285 | 24,285 | 24,285 | 24,285 | 24,285 | 24,285 |
| NYEARS | | 24 | | | | 24 |
| NFIRMS | | | 4,253 | | | ---------- |
| NINDS | | | | 363 | | 363 |
| NLOCS | | | | | 212 | 212 |
| RSQ | 0.073 | 0.079 | 0.446 | 0.114 | 0.088 | 0.132 |

**Panel B: MSA Fixed Effects on Cash Effective Tax Rates ($CETR_{it}$)**

| | **Model 1** | **Model 2** | **Model 3** | **Model 4** | **Model 5** | **Model 6** |
|---|---|---|---|---|---|---|
| Joint Significance (F-statistics) | | | | | | |
| $YEAR_t$ | | 10.84*** | | | | 10.39*** |
| $FIRM_i$ | | | 2.65*** | | | ---------- |
| $IND_i$ | | | | 2.88*** | | 2.77*** |
| $LOC_l$ (MSA) | | | | | 2.23*** | 2.01*** |
| Controls | YES | YES | YES | YES | YES | YES |
| N | 24,285 | 24,285 | 24,285 | 24,285 | 24,285 | 24,285 |
| NYEARS | | 24 | | | | 24 |
| NFIRMS | | | 4,253 | | | ---------- |
| NINDS | | | | 363 | | 363 |

| | | | | | | |
|---|---|---|---|---|---|---|
| NLOCS | | | | | 212 | 212 |
| RSQ | 0.077 | 0.086 | 0.409 | 0.115 | 0.094 | 0.139 |
| **Panel C: MSA Fixed Effects on FIN 48 Tax Reserve ($UTB_{it}$)** | | | | | | |
| | **Model 1** | **Model 2** | **Model 3** | **Model 4** | **Model 5** | **Model 6** |
| Joint Significance (F-statistics) | | | | | | |
| $YEAR_t$ | | 8.35*** | | | | 7.90*** |
| $FIRM_i$ | | | 17.75*** | | | ---------- |
| $IND_i$ | | | | 6.16*** | | 4.31*** |
| $LOC_l$ (MSA) | | | | | 14.16*** | 10.35*** |
| Controls | YES | YES | YES | YES | YES | YES |
| N | 16,478 | 16,478 | 16,478 | 16,478 | 16,478 | 16,478 |
| NYEARS | | 11 | | | | 11 |
| NFIRMS | | | 2,837 | | | ---------- |
| NINDS | | | | 318 | | 318 |
| NLOCS | | | | | 156 | 156 |
| RSQ | 0.101 | 0.105 | 0.808 | 0.198 | 0.207 | 0.274 |
| **Panel D: MSA Fixed Effects on 5-year Cash Effective Tax Rates ($CETR5_{it-4\sim t}$)** | | | | | | |
| | **Model 1** | **Model 2** | **Model 3** | **Model 4** | **Model 5** | **Model 6** |
| Joint Significance (F-statistics) | | | | | | |
| $YEAR_t$ | | 24.35*** | | | | 23.77*** |
| $FIRM_i$ | | | 10.15*** | | | ---------- |
| $IND_i$ | | | | 9.09*** | | 8.97*** |
| $LOC_l$ (MSA) | | | | | 5.47*** | 5.05*** |
| Controls | YES | YES | YES | YES | YES | YES |
| N | 17,553 | 17,553 | 17,553 | 17,553 | 17,553 | 17,553 |
| NYEARS | | 24 | | | | 24 |
| NFIRMS | | | 2,741 | | | ---------- |
| NINDS | | | | 347 | | 347 |

| | | | | | | |
|---|---|---|---|---|---|---|
| NLOCS | | | | | 162 | 162 |
| RSQ | 0.168 | 0.193 | 0.711 | 0.297 | 0.208 | 0.349 |

*, **, *** represent statistical significance at the 10%, 5%, and 1% levels, respectively.

This Table presents F-statistics testing the joint significance of the effects listed in the first column – YEAR, FIRM, IND, or LOC (MSA), and R-squared for fixed effect models. Panel A presents the results with ETR as the dependent variable, Panel B with CETR as the dependent variable, Panel C with UTB as the dependent variable, and Panel D with CETR5 as the dependent variable. Each column represents a regression nested within the first-stage model:

$$ETR_{it}(CETR_{it}/UTB_{it}/CETR5_{it-4\sim t}) = \alpha_0 + \Sigma_l \alpha_l LOC_l + \Sigma_t \alpha_t YEAR_t + \Sigma_i \alpha_i IND_i + \alpha_{str} STR_{it} + \Sigma_k \alpha_k CONTROL_{it}^k + \varepsilon_{it}$$

Model 1 includes only an intercept and the vector of time-varying firm-level controls. Models 2, 3, 4 and 5 are regressions, which include only one set of fixed effects (indicator variables) – YEAR, FIRM, IND, or LOC (MSA), but include no controls for the other effects (except the effects of the time-varying firm-level controls, which are included in all models). Model 6 includes all fixed effects except for firm.

## 4. Validity tests

Since we posit that the level of tax avoidance activities undertaken by a given firm is related to where the firm is located due to the influences from their geographical factors, it is expected that neighbouring MSAs should be relatively similar in their characteristics, and they should have similar impact on the firms' tax behaviour. We then test the relation between the distance from one MSA to another (i.e., a pair) and the absolute difference of each pair of the estimated MSA fixed effect coefficients found in Table 3 (see the main results). We define the geographical distance, $\text{Distance}_l$, as an indicator variable that equals to one if distance between a pair of MSA is greater than 125 kilometres; and zero if it is less or equal to 125 kilometres. $\text{DIFF}_{lETR}$, $\text{DIFF}_{lCETR}$ $\text{DIFF}_{lUTB}$ and $\text{DIFF}_{lCETR5}$ are the absolute differences of each pair of location fixed effects for ETR, CETR, UTB and CETR5, respectively. We regress $\text{DIFF}_l$ on $\text{Distance}_l$ with an intercept and robust standard errors. Among those MSAs that we use to estimate the effects of MSA fixed effects on MSA characteristics, 120 (or 86 or 77) MSAs yield 14,280 (or 7,310 or 5,852) pairs, excluding those pairs that have the same MSA in it. Table SA5 presents the results. It shows that the relationships between differences in location fixed effects and distances between MSAs are positively related for CETR, UTB and CETR5, indicating that the size of location fixed effects are more similar among MSAs that are geographically located closer. This finding affirms that the identified location fixed effects indeed show the geographical feature.

If our estimation on the location fixed effects is valid, we should also be able to observe location fixed effects being stable over time. We therefore, perform a sensitivity test that split the sample period into two using the mid-point of our sampling year. We use the year 2006 as the splitting point for the ETR, CETR and CETR5 samples and 2012 for the UTB samples. Reported in Table SA6, we find that in general, the location effects are significant in both sub-periods. There is no clear pattern indicating which period consistently dominates the results. Thus, consistent with our conjecture, the location-fixed effects appear to be stable over time.

**Table SA5: Differences in MSA Location Fixed Effects and Distances between MSAs**

| | $\text{DIFF}_l^{ETR}$ | $\text{DIFF}_l^{CETR}$ | $\text{DIFF}_l^{UTB}$ | $\text{DIFF}_l^{CETR5}$ |
|---|---|---|---|---|
| Intercept | 0.194*** | 0.213*** | 0.010*** | 0.090*** |
| | (10.89) | (11.75) | (9.43) | (8.26) |
| $\text{Distance}_l$ | 0.004 | 0.040** | 0.004*** | 0.025** |
| | (0.24) | (2.17) | (4.03) | (2.31) |
| N | 14,280 | 14,280 | 7,310 | 5,852 |
| RSQ | 0.0000 | 0.0002 | 0.0013 | 0.0009 |
| F-test | 0.06 | 4.70** | 16.24*** | 5.32** |

*, **, *** represent statistical significance at the 10%, 5%, and 1% levels, respectively.

This Table presents the results from testing the relation between the distance from one MSA to another (i.e. a pair) and the absolute difference of each pair of estimated MSA fixed effect coefficients with robust standard errors:

$DIFF_l^{ETR}(DIFF_l^{CETR}/DIFF_l^{UTB}/DIFF_l^{CETR5}) = \gamma_0 + \gamma_1 Distance_l$

$\text{DIFF}_l$ is the absolute difference of each pair of location fixed effects for ETR, CETR, UTB and CETR5, respectively. $\text{Distance}_l$ is an indicator variable that equals to one if distance between a pair of MSAs is greater than 125 kilometres; and zero if it is less or equal to 125 kilometres.

**Table SA6: Location Fixed Effects Over Time**

| | $\mathbf{ETR}_{it}$ | | $\mathbf{CETR}_{it}$ | | $\mathbf{UTB}_{it}$ | | $\mathbf{CETR5}_{it-4\sim t}$ | |
|---|---|---|---|---|---|---|---|---|
| | Pre 2006 | Post 2006 | Pre 2006 | Post 2006 | Pre 2012 | Post 2012 | Pre 2006 | Post 2006 |
| Joint Significance (F-statistics) | | | | | | | | |
| $\text{YEAR}_t$ | 2.79*** | 10.73*** | 9.48*** | 6.60*** | 3.83*** | 7.31*** | 20.36*** | 10.69*** |
| $\text{FIRM}_i$ | 2.88*** | 2.87*** | 2.04*** | 2.89*** | 15.34*** | 15.86*** | 11.01*** | 12.73*** |
| $\text{LOC}_l$ (MSA) | 1.71*** | 1.87*** | 1.58*** | 1.18 | 3.73*** | 2.24*** | 4.01*** | 4.63*** |
| Controls | YES | YES | YES | YES | YES | YES | YES | YES |
| N | 14,881 | 14,412 | 14,881 | 14,412 | 7,738 | 11,187 | 9,491 | 7,932 |
| NYEARS | 12 | 12 | 12 | 12 | 5 | 6 | 12 | 12 |
| NFIRMS | 4,037 | 3,030 | 4,037 | 3,030 | 2,289 | 2,788 | 2,348 | 2,067 |
| NLOCS | 201 | 165 | 201 | 165 | 139 | 151 | 148 | 132 |
| RSQ | 0.571 | 0.479 | 0.491 | 0.490 | 0.887 | 0.872 | 0.832 | 0.797 |

*, **, *** represent statistical significance at the 10%, 5%, and 1% levels, respectively.
This Table presents the results for the first half, i.e. pre 2006 (2012), and second half, i.e. post and include 2006 (2012) of the sample period. F-statistics test the joint significance of the effects listed in the first column – YEAR, FIRM, IND, or LOC (MSA); and R-squared for each regression.

## 5. Omitted variables tests

Prior studies find that governance factors such as board independence and CEO compensations are correlated with corporate tax avoidance (Armstrong et al., 2015; Gaertner, 2014). Thus, we additionally control for governance factors and CEO compensation to examine whether the geographic effects may work through the governance mechanisms. Specifically, we include measures of board size, percentage of independent directors, gender diversity on the board, CEO and Chair duality, and CFOs' board membership and CEO total compensation. These data are collected from Boardex and Execucomp. Including these variables severely reduces our sample size to a range between 11,369 (119) and 9,568 (107) firm-year (MSA) observations. Table SA7 reports the results. We find that controlling the above corporate governance and compensation factors in the reduced sample does not materially change our finding of significant location effects. As for these additional controls, we find that board independence exhibits a positive relationship with corporate tax avoidance.

Dyreng et al. (2010) find that executive fixed effects explain corporate tax avoidance. There is also evidence that CEOs prefer certain locations (Yonker, 2017). Therefore, we include CEO fixed effects in our models to test the robustness of our findings. We identify CEOs through the unique identifier of CEO from the ExecuComp database for the period from 1994 to 2017. Following Dyreng et al. (2010), each CEO in our sample is required to be employed by at least two different firms, for at least three years at each firm. This additional requirement significantly reduces our sample size. Table SA8 shows that CEO effect explains significant amount of variations in corporate tax avoidance, however, location fixed effects remain jointly significant in explaining tax avoidance across all full models after controlling for the CEO fixed effects. Overall, our findings suggest that the geographic effects are distinct from the executive effects.

**Table SA7: MSA Fixed Effects: Controlling for Corporate Governance and CEO Compensation**

**Panel A: MSA Fixed Effects on Effective Tax Rates ($ETR_{it}$)**

| | Model 1 | Model 2 | Model 3 | Model 4 | Model 5 | Model 6 |
|---|---|---|---|---|---|---|
| Joint Significance (F-statistics) | | | | | | |
| $YEAR_t$ | | 7.57*** | | | | 9.76*** |
| $FIRM_i$ | | | 2.80*** | | | 2.90*** |
| $IND_i$ | | | | 3.35*** | | ---------- |
| $LOC_l$ (MSA) | | | | | 1.87*** | 2.74*** |
| N | 11,369 | 11,369 | 11,369 | 11,369 | 11,369 | 11,369 |

| | | | | | | |
|---|---|---|---|---|---|---|
| NYEARS | | 19 | | | | 19 |
| NFIRMS | | | 1,623 | | | 1,623 |
| NINDS | | | | 300 | | ---------- |
| NLOCS | | | | | 119 | 119 |
| RSQ | 0.109 | 0.119 | 0.393 | 0.183 | 0.126 | 0.413 |
| Estimated Coefficients (t-tests): | | | | | | |
| $NUMDIRS_{it}$ | −0.000 | −0.001 | 0.003 | 0.007* | −0.000 | 0.004 |
| | (−0.10) | (−0.18) | (0.31) | (1.77) | (−0.10) | (0.50) |
| $PCTINDEP_{it}$ | −0.035*** | −0.024** | −0.052*** | −0.033*** | −0.039*** | −0.012 |
| | (−3.50) | (−2.25) | (−3.64) | (−3.11) | (−3.73) | (−0.77) |
| $GENDER_{it}$ | −0.012 | −0.005 | −0.010 | −0.023* | −0.012 | 0.005 |
| | (−0.97) | (−0.36) | (−0.51) | (−1.68) | (−0.90) | (0.23) |
| $CFOBOD_{it}$ | −0.000 | 0.003 | −0.003 | −0.001 | −0.000 | 0.003 |
| | (−0.13) | (0.84) | (−0.79) | (−0.37) | (−0.13) | (0.74) |
| $CEOCHAIR_{it}$ | 0.000 | 0.005 | −0.006* | −0.001 | −0.000 | −0.001 |
| | (0.04) | (1.53) | (−1.79) | (−0.40) | (−0.07) | (−0.27) |
| $CEOCOMP_{it}$ | 0.000 | −0.000 | 0.003 | 0.002 | 0.001 | 0.003 |
| | (0.15) | (−0.01) | (1.43) | (1.04) | (0.38) | (1.39) |
| Other Controls | YES | YES | YES | YES | YES | YES |

**Panel B: MSA Fixed Effects on Cash Effective Tax Rates ($CETR_{it}$)**

| | **Model 1** | **Model 2** | **Model 3** | **Model 4** | **Model 5** | **Model 6** |
|---|---|---|---|---|---|---|
| Joint Significance (F-statistics) | | | | | | |
| $YEAR_t$ | | 5.16*** | | | | 9.44*** |
| $FIRM_i$ | | | 3.25*** | | | 3.32*** |
| $IND_i$ | | | | 3.13*** | | ---------- |
| $LOC_l$ (MSA) | | | | | 2.33*** | 2.12*** |
| N | 11,369 | 11,369 | 11,369 | 11,369 | 11,369 | 11,369 |

| | | | | | | |
|---|---|---|---|---|---|---|
| NYEARS | | 19 | | | | 19 |
| NFIRMS | | | 1,623 | | | 1,623 |
| NINDS | | | | 300 | | ---------- |
| NLOCS | | | | | 119 | 119 |
| RSQ | 0.112 | 0.120 | 0.424 | 0.182 | 0.134 | 0.441 |
| Estimated Coefficients (t-tests): | | | | | | |
| $NUMDIRS_{it}$ | −0.002 | −0.002 | 0.009 | −0.001 | −0.001 | 0.013 |
| | (−0.53) | (−0.40) | (0.88) | (−0.21) | (−0.27) | (1.29) |
| $PCTINDEP_{it}$ | −0.033*** | −0.041*** | −0.027 | −0.031** | −0.038*** | −0.018 |
| | (−2.71) | (−3.15) | (−1.59) | (−2.38) | (−3.05) | (−0.91) |
| $GENDER_{it}$ | −0.037** | −0.039** | 0.006 | −0.007 | −0.034** | −0.044* |
| | (−2.43) | (−2.52) | (0.24) | (−0.40) | (−2.16) | (−1.72) |
| $CFOBOD_{it}$ | 0.009** | 0.004 | 0.002 | 0.009** | 0.010*** | −0.003 |
| | (2.43) | (1.09) | (0.43) | (2.54) | (2.85) | (−0.57) |
| $CEOCHAIR_{it}$ | −0.001 | 0.003 | −0.007* | 0.000 | −0.002 | 0.006 |
| | (−0.20) | (0.64) | (−1.81) | (0.09) | (−0.59) | (1.31) |
| $CEOCOMP_{it}$ | −0.005** | −0.005** | −0.005* | −0.007*** | −0.006*** | −0.002 |
| | (−2.49) | (−2.39) | (−1.95) | (−3.53) | (−2.84) | (−0.72) |
| Other Controls | YES | YES | YES | YES | YES | YES |

**Panel C: MSA Fixed Effects on FIN 48 Tax Reserve ($UTB_{it}$)**

| | **Model 1** | **Model 2** | **Model 3** | **Model 4** | **Model 5** | **Model 6** |
|---|---|---|---|---|---|---|
| Joint Significance (F-statistics) | | | | | | |
| $YEAR_t$ | | 11.23*** | | | | 8.14*** |
| $FIRM_i$ | | | 15.35*** | | | 14.23*** |
| $IND_i$ | | | | 5.49*** | | ---------- |
| $LOC_l$ (MSA) | | | | | 9.21*** | 4.52*** |
| N | 9,568 | 9,568 | 9,568 | 9,568 | 9,568 | 9,568 |

| | | | | | | |
|---|---|---|---|---|---|---|
| NYEARS | | 11 | | | | 11 |
| NFIRMS | | | 1,395 | | | 1,395 |
| NINDS | | | | 283 | | ---------- |
| NLOCS | | | | | 112 | 112 |
| RSQ | 0.223 | 0.232 | 0.786 | 0.334 | 0.299 | 0.794 |
| Estimated Coefficients (t-tests): | | | | | | |
| $NUMDIRS_{it}$ | 0.001 | 0.000 | 0.001* | 0.001** | 0.001* | 0.001 |
| | (1.47) | (0.89) | (1.85) | (2.31) | (1.71) | (1.36) |
| $PCTINDEP_{it}$ | 0.008*** | 0.007*** | 0.002 | 0.008*** | 0.010*** | 0.002 |
| | (4.93) | (4.18) | (1.37) | (4.71) | (6.38) | (1.25) |
| $GENDER_{it}$ | 0.008*** | 0.005*** | 0.006*** | 0.009*** | 0.008*** | 0.003* |
| | (5.07) | (3.26) | (3.34) | (5.23) | (5.33) | (1.78) |
| $CFOBOD_{it}$ | 0.001** | 0.001** | 0.000 | 0.000 | 0.000 | 0.000 |
| | (2.52) | (2.05) | (1.02) | (1.10) | (0.61) | (1.03) |
| $CEOCHAIR_{it}$ | −0.002*** | −0.002*** | −0.001* | −0.002*** | −0.002*** | −0.000 |
| | (−6.80) | (−5.46) | (−1.90) | (−5.77) | (−6.10) | (−1.54) |
| $CEOCOMP_{it}$ | −0.000 | 0.000 | 0.000 | 0.000 | −0.000 | 0.000 |
| | (−0.38) | (0.28) | (0.28) | (0.00) | (−0.54) | (1.35) |
| Other Controls | YES | YES | YES | YES | YES | YES |

**Panel D: MSA Fixed Effects on 5-Year Cash ETR ($CETR5_{it-4\sim t}$)**

| | **Model 1** | **Model 2** | **Model 3** | **Model 4** | **Model 5** | **Model 6** |
|---|---|---|---|---|---|---|
| Joint Significance (F-statistics) | | | | | | |
| $YEAR_t$ | | 6.72*** | | | | 13.31*** |
| $FIRM_i$ | | | 11.02*** | | | 11.35*** |
| $IND_i$ | | | | 7.63*** | | ---------- |
| $LOC_l$ (MSA) | | | | | 4.87*** | 5.77*** |
| N | 10,022 | 10,022 | 10,022 | 10,022 | 10,022 | 10,022 |

| | | | | | | |
|---|---|---|---|---|---|---|
| NYEARS | | 19 | | | | 19 |
| NFIRMS | | | 1,385 | | | 1,385 |
| NINDS | | | | 290 | | ---------- |
| NLOCS | | | | | 107 | 107 |
| RSQ | 0.198 | 0.207 | 0.710 | 0.346 | 0.237 | 0.727 |
| Estimated Coefficients (t-tests): | | | | | | |
| $\text{NUMDIRS}_{it}$ | 0.004 | 0.004 | 0.005 | 0.006* | 0.004 | 0.008* |
| | (1.40) | (1.25) | (1.06) | (1.92) | (1.42) | (1.67) |
| $\text{PCTINDEP}_{it}$ | −0.040*** | −0.018** | −0.044*** | −0.041*** | −0.043*** | 0.009 |
| | (−4.71) | (−1.98) | (−5.06) | (−4.87) | (−5.06) | (0.89) |
| $\text{GENDER}_{it}$ | −0.043*** | −0.038*** | 0.036*** | −0.001 | −0.035*** | 0.015 |
| | (−4.13) | (−3.65) | (2.85) | (−0.07) | (−3.34) | (1.17) |
| $\text{CFOBOD}_{it}$ | 0.004 | 0.001 | 0.002 | 0.004* | 0.007*** | −0.000 |
| | (1.58) | (0.55) | (1.00) | (1.82) | (2.79) | (−0.03) |
| $\text{CEOCHAIR}_{it}$ | 0.005** | −0.001 | 0.002 | 0.007*** | 0.004 | 0.001 |
| | (2.15) | (−0.28) | (1.08) | (3.03) | (1.54) | (0.36) |
| $\text{CEOCOMP}_{it}$ | −0.002 | −0.002 | −0.003** | −0.004*** | −0.002* | −0.001 |
| | (−1.08) | (−1.37) | (−2.02) | (3.21) | (−1.74) | (−1.13) |
| Other Controls | YES | YES | YES | YES | YES | YES |

*, **, *** represent statistical significance at the 10%, 5%, and 1% levels, respectively.

This Table presents results after controlling for corporate governance factors and CEO compensation.

Panel A presents the results with ETR as the dependent variable, Panel B with CETR as the dependent variable, Panel C with UTB as the dependent variable, and Panel D with CETR5 as the dependent variable.

Each column represents a regression nested within the first-stage model:

$$ETR_{it}(CETR_{it}/UTB_{it}/CETR5_{it-4\sim t}) = \alpha_0 + \Sigma_l \alpha_l LOC_l + \Sigma_t \alpha_t YEAR_t + \Sigma_i \alpha_i FIRM_i + \alpha_{str} STR_{it} + \Sigma_k \alpha_k CONTROL_{it}^k + \varepsilon_{it}$$

Model 1 includes only an intercept and the vector of time-varying firm-level controls. Models 2, 3, 4 and 5 are regressions, which include only one set of fixed effects (indicator variables) – YEAR, FIRM, IND, or LOC (MSA), but include no controls for the other effects (except the effects of the time-varying firm-level controls, which are included in all models). The final model, Model 6, is a test of each set of effects in the presence of the other effects, except for industry.

**Table SA8: MSA Fixed Effects: Controlling for Executive Fixed Effects**

| **Panel A: MSA Fixed Effects on Effective Tax Rates ($ETR_{it}$)** | | | | | | | |
|---|---|---|---|---|---|---|---|
| | **Model 1** | **Model 2** | **Model 3** | **Model 4** | **Model 5** | **Model 6** | **Model 7** |
| Joint Significance (F-statistics) | | | | | | | |
| $YEAR_t$ | | 6.17*** | | | | | 5.04*** |
| $FIRM_i$ | | | 2.52*** | | | | ---------- |
| $IND_i$ | | | | 2.94*** | | | 3.55*** |
| $LOC_l$ (MSA) | | | | | 1.79*** | | 2.44*** |
| $CEO_i$ | | | | | | 2.45*** | 2.35*** |
| Controls | YES | YES | YES | YES | YES | YES | YES |
| N | 12,289 | 12,289 | 12,289 | 12,289 | 12,289 | 12,289 | 12,289 |
| NYEARS | | 24 | | | | | 24 |
| NFIRMS | | | 1,676 | | | | ---------- |
| NINDS | | | | 306 | | | 306 |
| NLOCS | | | | | 172 | | 172 |
| NCEO | | | | | | 2,807 | 2,807 |
| RSQ | 0.087 | 0.097 | 0.348 | 0.151 | 0.110 | 0.471 | 0.490 |
| **Panel B: MSA Fixed Effects on Cash Effective Tax Rates ($CETR_{it}$)** | | | | | | | |
| | **Model 1** | **Model 2** | **Model 3** | **Model 4** | **Model 5** | **Model 6** | **Model 7** |
| Joint Significance (F-statistics) | | | | | | | |
| $YEAR_t$ | | 4.99*** | | | | | 5.86*** |
| $FIRM_i$ | | | 2.60*** | | | | ---------- |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| $IND_i$ | | | | 2.65*** | | | 1.92** |
| $LOC_l$ (MSA) | | | | | 1.95*** | | 2.30*** |
| $CEO_i$ | | | | | | 2.49*** | 2.40*** |
| Controls | YES | YES | YES | YES | YES | YES | YES |
| N | 12,289 | 12,289 | 12,289 | 12,289 | 12,289 | 12,289 | 12,289 |
| NYEARS | | 24 | | | | | 24 |
| NFIRMS | | | 1,676 | | | | ---------- |
| NINDS | | | | 306 | | | 306 |
| NLOCS | | | | | 172 | | 172 |
| NCEO | | | | | | 2,807 | 2,807 |
| RSQ | 0.088 | 0.096 | 0.354 | 0.145 | 0.112 | 0.475 | 0.493 |

**Panel C: MSA Fixed Effects on FIN 48 Tax Reserve ($UTB_{it}$)**

| | **Model 1** | **Model 2** | **Model 3** | **Model 4** | **Model 5** | **Model 6** | **Model 7** |
|---|---|---|---|---|---|---|---|
| Joint Significance (F-statistics) | | | | | | | |
| $YEAR_t$ | | 9.03*** | | | | | 4.34*** |
| $FIRM_i$ | | | 16.34*** | | | | ---------- |
| $IND_i$ | | | | 5.57*** | | | 7.01*** |
| $LOC_l$ (MSA) | | | | | 5.71*** | | 3.95*** |
| $CEO_i$ | | | | | | 14.92*** | 14.04*** |
| Controls | YES | YES | YES | YES | YES | YES | YES |
| N | 7,399 | 7,399 | 7,399 | 7,399 | 7,399 | 7,399 | 7,399 |
| NYEARS | | 11 | | | | | 11 |
| NFIRMS | | | 1,144 | | | | ---------- |
| NINDS | | | | 269 | | | 269 |
| NLOCS | | | | | 137 | | 137 |
| NCEO | | | | | | 1,539 | 1,539 |
| RSQ | 0.181 | 0.191 | 0.795 | 0.301 | 0.260 | 0.834 | 0.841 |

**Panel D: MSA Fixed Effects on 5-year Cash Effective Tax Rates ($CETR5_{it-4\sim t}$)**

| | Model 1 | Model 2 | Model 3 | Model 4 | Model 5 | Model 6 | Model 7 |
|---|---|---|---|---|---|---|---|
| Joint Significance (F-statistics) | | | | | | | |
| $YEAR_t$ | | 16.26*** | | | | | 20.86*** |
| $FIRM_i$ | | | 9.89*** | | | | ---------- |
| $IND_i$ | | | | 7.26*** | | | 8.51*** |
| $LOC_l$ (MSA) | | | | | 5.42*** | | 5.15*** |
| $CEO_i$ | | | | | | 10.74*** | 9.45*** |
| Controls | YES | YES | YES | YES | YES | YES | YES |
| N | 10,593 | 10,593 | 10,593 | 10,593 | 10,593 | 10,593 | 10,593 |
| NYEARS | | 24 | | | | | 24 |
| NFIRMS | | | 1,444 | | | | ---------- |
| NINDS | | | | 295 | | | 295 |
| NLOCS | | | | | 141 | | 141 |
| NCEO | | | | | | 2,329 | 2,329 |
| RSQ | 0.203 | 0.229 | 0.689 | 0.340 | 0.257 | 0.802 | 0.821 |

*, **, *** represent statistical significance at the 10%, 5%, and 1% levels, respectively.

This Table presents F-statistics testing the joint significance of the effects listed in the first column – YEAR, FIRM, IND, LOC (MSA), or CEO; and R-squared for fixed effect models. Panel A presents the results with ETR as the dependent variable, Panel B with CETR as the dependent variable, Panel C with UTB as the dependent variable, and Panel D with CETR5 as the dependent variable. Each column represents a regression nested within the first-stage model:

$$ETR_{it}(CETR_{it}/UTB_{it}/CETR5_{it-1\sim t}) = \alpha_0 + \Sigma_l \alpha_l LOC_l + \Sigma_t \alpha_t YEAR_t + \Sigma_i \alpha_i FIRM_i + \Sigma_m \alpha_m CEO_m + \alpha_{str} STR_{it} + \Sigma_k \alpha_k CONTROL_{it}^k + \varepsilon_{it}$$

Model 1 includes only an intercept and the vector of time-varying firm-level controls. Models 2, 3, 4, 5, and 6 are regressions, which include only one set of fixed effects (indicator variables) – YEAR, FIRM, IND, LOC (MSA), or CEO, but include no controls for the other effects (except the effects of the time-varying firm-level controls, which are included in all models). The final model, Model 7, is a test of each set of effects in the presence of the other effects, except for industry.