

Supplementary Material

Tax avoidance and audit quality: evidence from Spanish listed and non-listed companies

Appendices

Appendix 1. Variable definitions

Variables	Definitions
Dependent Variable (measured as TA)	
<i>MCASHETR</i>	Modified Cash Effective Tax Rate, measured as the ratio of total cash taxes paid to pre-tax income, scaled by Statutory Tax Rate
<i>BTG</i>	Book Tax Gap, calculated as the natural logarithm of the absolute value of the total book-tax differences (pre-tax profit minus estimated taxable income)
<i>BTD</i>	Book Tax Differences, defined as pre-tax profit minus estimated taxable income, scaled by previous year total assets.
<i>DTAX</i>	Discretionary Tax Differences, computed following Frank et al. (2009) and Monterrey Mayoral & Sánchez Segura (2022).
Main Variables	
<i>BIG4</i>	Auditor size: a dummy that takes the value of 1 if the auditor is a BIG4; and 0 otherwise
<i>TENURE</i>	Auditor experience: natural logarithm of the number of consecutive years audited by the same auditor
Control Variables	
<i>SIZE</i>	Natural logarithm of total assets.
<i>LEV</i>	Total leverage scaled by total assets.
<i>INTINT</i>	Total intangible assets scaled by total assets.
<i>PPEINT</i>	Book value of gross property, plant, and equipment scaled by total assets.
<i>INVINT</i>	Inventories scaled by total assets.
<i>ROA</i>	Pretax income scaled by total assets.
<i>SGROW</i>	Firm's sales growth, measured as the percentage change in sales between year t and t-1.
<i>FOREIGN</i>	A dummy that takes the value of 1 if the firm carries out activities in foreign countries; and 0 otherwise.
<i>FORAL</i>	A dummy that takes the value of 1 if the firm is domiciled in a Foral territory (Basque Country or Navarre); and 0 otherwise.
<i>CANARIAS</i>	A dummy that takes the value of 1 if the firm is domiciled in the Canary Islands; and 0 otherwise.
<i>TRLAC11</i>	A dummy that takes the value of 1 if 2011 Spanish Auditing Law (<i>Texto Refundido de la Ley de Auditoría de Cuentas de 2011</i>) is applied; and 0 otherwise.
<i>LAC15</i>	A dummy that takes the value of 1 if 2015 Spanish Auditing Law (<i>Ley de Auditoría de Cuentas de 2015</i>) is applied; and 0 otherwise.
<i>GDP</i>	Rate of year-on-year variation in GDP

Appendix 2. Descriptive statistics and correlations

This table reports the descriptive statistics for all variables used in the model (1). In Panel A, SD is standard deviation, and P25 and P75 are the first and the third quartile, respectively. Panel B shows the Pearson correlation coefficients. See Appendix 1 for detailed variable definitions. ***, **, and * denote significance at 1%, 5%, and 10%, respectively.

Panel A. Descriptive statistics

Variable	Mean			SD			P25			Median			P75		
	All	Non-listed	Listed	All	Non-listed	Listed	All	Non-listed	Listed	All	Non-listed	Listed	All	Non-listed	Listed
<i>MCASHETR</i>	0.886	0.887	0.563	0.543	0.543	0.479	0.549	0.552	0.166	0.889	0.890	0.443	1.080	1.082	0.891
<i>BTG</i>	3.443	3.420	8.599	4.065	4.054	3.016	2.048	2.084	6.534	4.008	4.000	9.343	5.847	5.825	10.635
<i>BIG4</i>	0.263	0.261	0.846	0.441	0.439	0.361	0.000	0.000	1.000	0.000	0.000	1.000	1.000	1.000	1.000
<i>TENURE</i>	1.821	1.822	1.584	0.795	0.795	0.804	1.386	1.386	1.099	1.946	1.946	1.792	2.398	2.398	2.197
<i>SIZE</i>	4.266	4.262	5.294	0.550	0.545	0.687	3.884	3.883	4.730	4.159	4.156	5.048	4.536	4.530	5.779
<i>LEV</i>	0.501	0.501	0.476	0.231	0.232	0.226	0.315	0.316	0.346	0.505	0.505	0.506	0.683	0.683	0.630
<i>INTINT</i>	0.028	0.027	0.049	0.087	0.087	0.084	0.001	0.001	0.001	0.003	0.003	0.008	0.013	0.012	0.061
<i>PPEINT</i>	0.219	0.220	0.148	0.219	0.203	0.196	0.054	0.054	0.014	0.167	0.168	0.096	0.329	0.329	0.184
<i>INVINT</i>	0.179	0.180	0.124	0.159	0.159	0.135	0.045	0.045	0.027	0.146	0.146	0.085	0.272	0.273	0.152
<i>ROA</i>	5.533	5.520	8.256	5.249	5.240	6.565	1.685	1.682	2.652	4.004	3.994	7.120	7.674	7.654	11.359
<i>SGROW</i>	0.076	0.075	0.189	1.257	1.258	1.033	-0.042	-0.043	-0.027	0.031	0.031	0.047	0.114	0.114	0.123
<i>FOREIGN</i>	0.694	0.694	0.646	0.461	0.461	0.480	0.000	0.000	0.000	1.000	1.000	1.000	1.000	1.000	1.000
<i>FORAL</i>	0.074	0.073	0.254	0.262	0.261	0.436	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000
<i>CANARIAS</i>	0.031	0.031	0.000	0.172	0.173	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
<i>TRLAC11</i>	0.662	0.662	0.645	0.473	0.473	0.480	0.000	0.000	0.000	1.000	1.000	1.000	1.000	1.000	1.000
<i>LAC15</i>	0.324	0.324	0.312	0.468	0.468	0.465	0.000	0.000	0.000	0.000	0.000	0.000	1.000	1.000	1.000
<i>GDP</i>	0.673	0.674	0.578	2.341	2.341	2.358	-0.811	-0.810	-1.440	0.892	0.890	0.890	2.973	2.970	2.970

Panel B. Correlation matrix

Variable	<i>MCASHETR</i>	<i>BTG</i>	<i>BIG4</i>	<i>TENURE</i>	<i>SIZE</i>	<i>LEV</i>	<i>INTINT</i>	<i>PPEINT</i>	<i>INVINT</i>	<i>ROA</i>	<i>SGROW</i>	<i>FOREIGN</i>	<i>FORAL</i>	<i>CANARIAS</i>	<i>TRLAC11</i>	<i>LAC15</i>	<i>GDP</i>
<i>MCASHETR</i>	1.000																
<i>BTG</i>	0.129***	1.000															
<i>BIG4</i>	0.014**	0.249***	1.000														
<i>TENURE</i>	0.005	0.063***	0.076	1.000													
<i>SIZE</i>	-0.055***	0.485***	0.427***	0.067***	1.000												
<i>LEV</i>	0.006	-0.089***	0.093***	-0.059***	0.026***	1.000											
<i>INTINT</i>	0.008	0.062***	0.212***	-0.017***	0.122***	0.058***	1.000										
<i>PPEINT</i>	-0.025***	0.042***	-0.006	-0.029***	0.050***	-0.036***	-0.103***	1.000									
<i>INVINT</i>	0.012**	-0.155***	-0.161***	0.023***	-0.133***	0.094***	-0.150***	-0.247***	1.000								
<i>ROA</i>	-0.114***	0.189***	0.153***	-0.005	0.036***	-0.239***	0.013**	-0.059***	-0.160***	1.000							
<i>SGROW</i>	-0.049***	0.029***	0.017***	-0.007	0.033***	0.012**	-0.002	-0.017***	0.014***	0.023***	1.000						
<i>FOREIGN</i>	-0.018***	0.060***	-0.061***	0.034***	0.034***	-0.056***	-0.133***	-0.138***	0.149***	0.012**	-0.016***	1.000					
<i>FORAL</i>	-0.052***	0.137***	0.047***	-0.010**	0.044***	-0.013***	-0.020***	-0.009*	0.005	-0.006	0.006	0.048***	1.000				
<i>CANARIAS</i>	-0.095***	0.101***	0.012**	0.006	0.037***	-0.102***	-0.024***	0.097***	-0.076***	0.002	-0.002	-0.127***	-0.050***	1.000			
<i>TRLAC11</i>	0.020***	0.107***	0.019***	0.293***	0.051***	-0.099***	0.017***	-0.042***	0.041***	0.017***	0.001	0.001	-0.000	0.003	1.000		
<i>LAC15</i>	0.012**	0.111***	0.018***	0.229***	0.071***	-0.073***	-0.012**	-0.032***	0.052***	0.070***	0.001	-0.002	0.005	0.000	0.495***	1.000	
<i>GDP</i>	0.004	0.142***	0.017***	0.197***	0.051***	-0.064***	-0.004	-0.032***	0.052***	0.066***	0.016***	-0.000	0.006	0.004	0.462***	0.582***	1.000

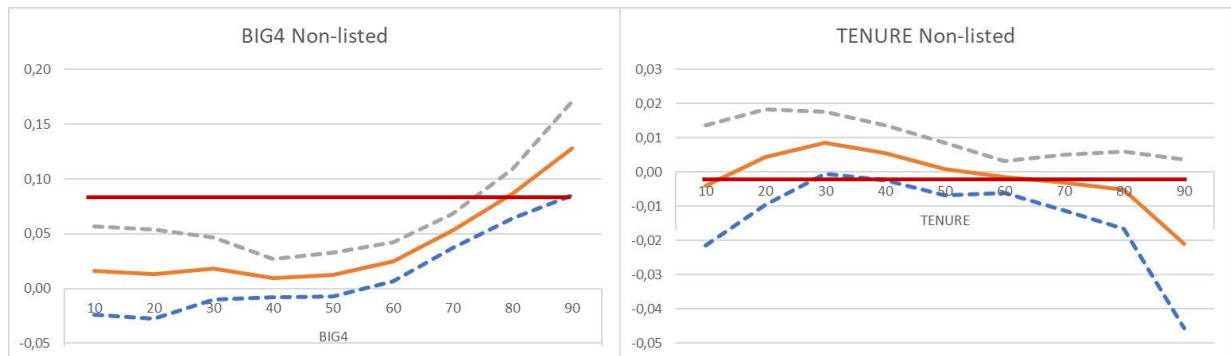
Appendix 3. Wald tests of coefficient equality across quantiles

This table reports F-statistics from Wald tests on simultaneous quantile regression coefficients (Tables III-IV) with 1,000 bootstrap replications: (1) General non-linearity test (all quantiles); (2) Critical tests: lower vs. upper quantiles; (3) Tests against the median. See Appendix 1 for detailed variable definitions. ***, **, and * denote significance at 1%, 5%, and 10%, respectively.

Test Wald		TA = MCASHETR		TA = BTG	
		<i>BIG4</i>	<i>TENURE</i>	<i>BIG4</i>	<i>TENURE</i>
		F value	F value	F value	F value
Non-listed		Observations: 18,974		Observations: 25,999	
1	Q(0.10)=Q(0.20)=...=Q(0.90)	(5.93)***	(0.97)	(3.94)***	(2.62)***
2	Q(0.90) = Q(0.10)	(18.59)***	(1.37)	(12.52)***	(11.04)***
2	Q(0.80) = Q(0.20)	(17.85)***	(1.20)	(13.88)***	(2.87)*
2	Q(0.70) = Q(0.30)	(8.15)***	(3.70)**	(16.34)***	(4.03)**
3	Q(0.10) = Q(0.50)	(0.05)	(0.36)	(5.03)**	(7.80)***
3	Q(0.20) = Q(0.50)	(0.00)	(0.25)	(6.28)***	(0.34)
3	Q(0.80) = Q(0.50)	(34.39)***	(1.19)	(9.55)***	(6.60)***
3	Q(0.90) = Q(0.50)	(27.34)***	(3.14)*	(18.38)***	(5.66)**
Listed		Observations: 109		Observations: 133	
1	Q(0.10)=Q(0.20)=...=Q(0.90)	(0.16)	(0.85)	(0.49)	(0.16)
2	Q(0.90) = Q(0.10)	(0.90)	(1.38)	(0.25)	(0.14)
2	Q(0.80) = Q(0.20)	(0.38)	(0.03)	(0.20)	(0.00)
2	Q(0.70) = Q(0.30)	(0.16)	(0.10)	(0.21)	(0.13)
3	Q(0.10) = Q(0.50)	(0.16)	(0.77)	(0.19)	(0.36)
3	Q(0.20) = Q(0.50)	(0.27)	(0.68)	(0.17)	(0.00)
3	Q(0.80) = Q(0.50)	(0.07)	(0.10)	(0.02)	(0.02)
3	Q(0.90) = Q(0.50)	(0.62)	(3.37)*	(0.06)	(0.14)

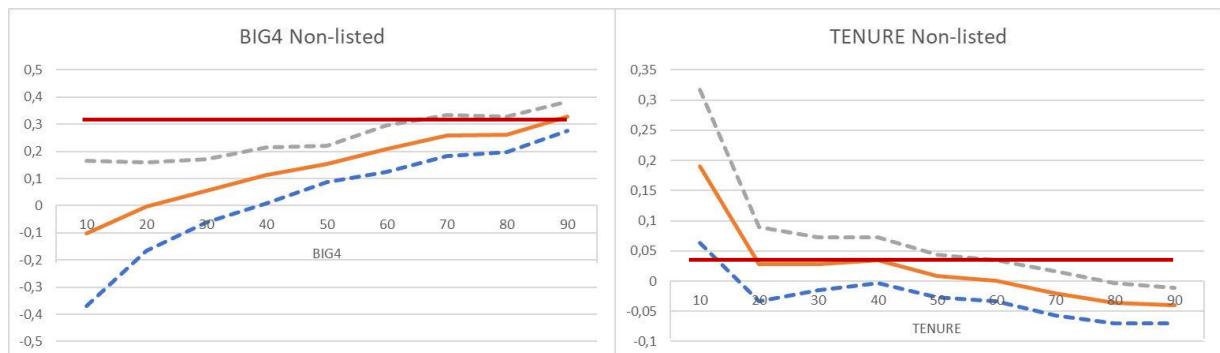
Appendix 4. Figures

Figure 1. Quantile Regression Estimates (MCASHETR) - Non-listed companies



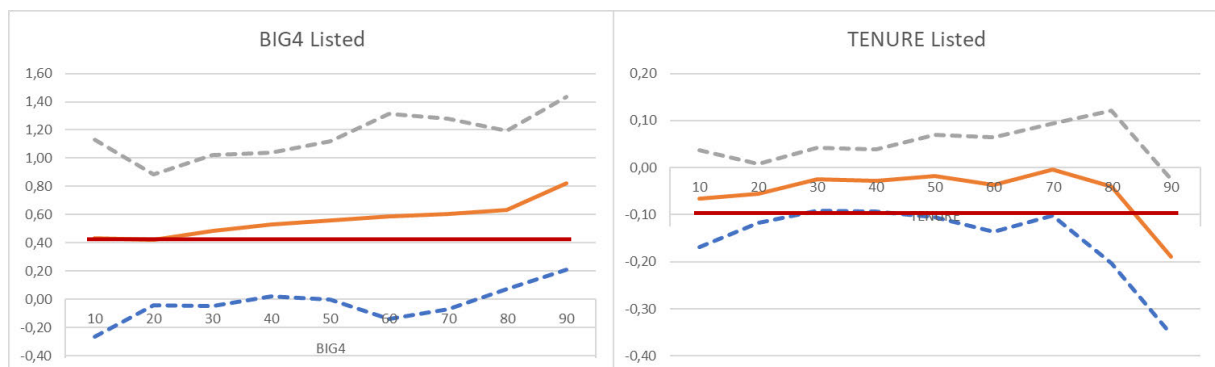
The orange solid line plots quantile estimates, the dashed lines the 95% pointwise confidence interval and the red solid line the PCSE estimate.

Figure 2. Quantile Regression Estimates (BTG) - Non-listed companies



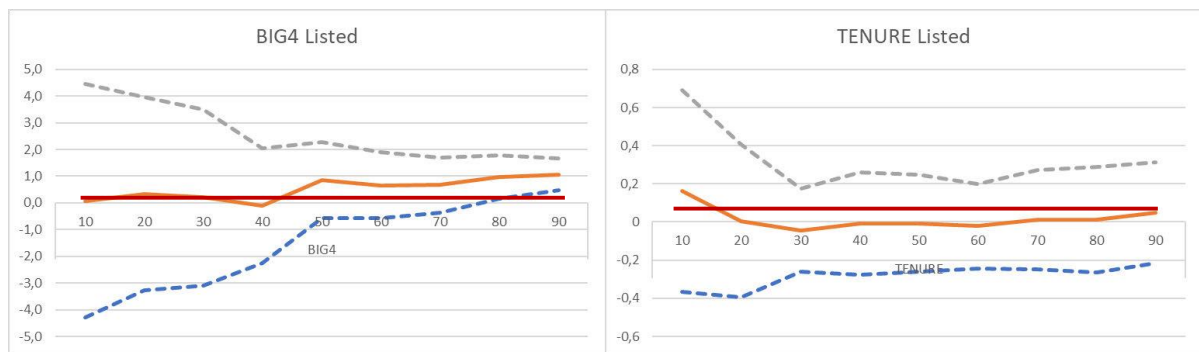
Same colour coding as in Figure 1.

Figure 3. Quantile Regression Estimates (MCASHETR) - Listed companies



Same colour coding as in Figure 1.

Figure 4. Quantile Regression Estimates (BTG) - Listed companies



Same colour coding as in Figure 1.

Appendix 5. Robustness Tests

Alternative Measures of Tax Avoidance

Given that there is no clear variable for measuring tax avoidance, in this section we employ two additional measures of TA that have been widely used in the previous literature. First, we use a second book-tax gap measure, BTG, originally proposed by Manzon & Plesko (2002), and subsequently applied in various studies, including Monterrey Mayoral & Sánchez Segura (2010) for the Spanish context. This measure is calculated by dividing the difference between pre-tax net income and estimated taxable income by total assets from the previous year. Second, we employ a measure based on the approach developed by Frank et al. (2009), which has also been extensively used in the literature, with different adaptations, such as that of Monterrey Mayoral & Sánchez Segura (2022). This variable, known as DTAX, is obtained from the residuals of a regression model in which the dependent variable is the book-tax gap. The independent variables vary across studies, although the original model by Frank et al. (2009) aimed to capture the main permanent differences defined under U.S. tax law. As Monterrey Mayoral & Sánchez Segura (2022) point out, these variables must be adjusted to the Spanish context. However, the lack of data on these permanent differences means that all variables used in the model are proxies, which may complicate the estimation.

It is widely acknowledged that all measures of TA are proxies. Therefore, researchers tend to employ multiple indicators to enhance the robustness and validity of their findings. According to Frank et al. (2009), ETR-based measures, which involve dividing tax expense or cash taxes paid by pre-tax accounting income, have the drawback of relying on a potentially distorted denominator. These authors argue that firms engaging in aggressive tax strategies may simultaneously adopt aggressive accounting practices, thus altering reported earnings and affecting the ETR calculation. Nonetheless, in our view, when earnings are manipulated, balance sheet figures such as total assets may also be distorted, meaning that all proxies are subject to limitations. This reinforces the importance of using multiple TA measures in order to improve the reliability and consistency of the results.

Table A1 presents the model estimations using both PCSE and quantile regressions for the two alternative measures of TA, BTG and DTAX. In addition, Table A2 presents the Wald tests. Because BTG and DTAX retain the sign of the book-tax difference, higher values are generally interpreted as reflecting greater tax avoidance, whereas the absolute-value BTG measure captures the magnitude but not the direction of the gap. Accordingly, these alternative signed measures complement the main analysis by assessing whether the documented distributional heterogeneity persists when the direction of the book-tax difference is retained.

For BIG4 in non-listed firms (Table A1, Panels A–B), the results are broadly consistent with those obtained using BTG (Table III), revealing heterogeneous associations across the conditional distributions of BTG and DTAX. In the upper quantiles, BIG4 is positively and statistically significantly associated with both measures, starting from quantile 0.6 for DTAX and quantile 0.8 for BTG. These findings are consistent with those reported by Chen et al. (2023) and Monterrey Mayoral and Sánchez Segura (2022) and provide additional evidence that, among non-listed firms, the association between BIG4 and TA varies across the distribution.

Table A1. PCSE and Quantile Regression Estimations using Alternative Measures of TA

This table reports the results of model (1) estimated using PCSE and quantile regression across nine quantiles of the conditional TA distribution. In this case, TA is measured using BTD and DTAX. See Appendix 1 for detailed variable definitions. ***, **, and * denote significance at 1%, 5%, and 10%, respectively.

Panel A. Estimation Results using BTD to measure TA - Non-listed

Observations: 24,776	PCSE	Quantiles								
		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
<i>BIG4</i>	-0.0003 (-0.30)	-0.0040 (-9.17)***	-0.0009 (-6.56)***	-0.0004 (-7.11)***	-0.0008 (-7.68)***	-0.0009 (-5.57)***	-0.0005 (-1.49)	0.0008 (1.62)	0.0027 (4.14)***	0.0074 (6.37)***
<i>TENURE</i>	-0.0001 (-0.35)	0.0001 (1.45)	0.0001 (2.57)***	0.0000 (2.22)**	0.0001 (2.07)**	0.0001 (1.29)	-0.0000 (-0.33)	-0.0001 (-0.66)	-0.0002 (-1.17)	-0.0003 (-0.92)
Pseudo R ²	0.1280	0.0221	0.0045	0.0066	0.0233	0.0496	0.0782	0.1137	0.1657	0.2437

Panel B. Estimation Results using DTAX to measure TA - Non-listed

Observations: 16,747	PCSE	Quantiles								
		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
<i>BIG4</i>	-0.0001 (-0.12)	-0.0019 (-5.92)***	-0.0007 (-5.99)***	-0.0009 (-6.07)***	-0.0006 (-2.83)***	0.0001 (0.24)	0.0011 (2.28)**	0.0024 (4.15)***	0.0044 (4.66)***	0.0092 (6.99)***
<i>TENURE</i>	0.0005 (1.35)	0.0003 (2.85)***	0.0002 (3.22)***	0.0002 (2.39)**	0.0001 (1.67)*	0.0000 (0.52)	-0.0000 (-0.49)	-0.0002 (-1.45)	-0.0003 (-1.52)	-0.0012 (-3.03)***
Pseudo R ²	0.1471	0.0093	0.0079	0.0231	0.0452	0.0709	0.1021	0.1428	0.1962	0.2736

Panel C. Estimation Results using BTD to measure TA - Listed

Observations: 124	PCSE	Quantiles								
		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
<i>BIG4</i>	0.0055 (0.31)	0.0139 (0.27)	0.0309 (0.56)	0.0181 (0.31)	0.0260 (0.41)	0.0089 (0.13)	0.0058 (0.08)	0.0085 (0.10)	0.0261 (0.27)	0.0219 (0.18)
<i>TENURE</i>	-0.0046 (-1.13)	-0.0143 (-1.47)	-0.0095 (-1.13)	-0.0036 (-0.49)	-0.0054 (-0.80)	-0.0046 (-0.67)	-0.0045 (-0.64)	0.0026 (0.36)	0.0039 (0.50)	0.0085 (0.87)
Pseudo R ²	0.0000	0.3937	0.4026	0.4472	0.4757	0.5059	0.5332	0.5624	0.5929	0.6233

Panel D. Estimation Results using DTAX to measure TA - Listed

Observations: 107	PCSE	Quantiles								
		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
<i>BIG4</i>	-0.0098 (-0.45)	-0.0216 (-0.35)	0.0274 (0.45)	0.0353 (0.60)	0.0488 (0.88)	0.0161 (0.32)	0.0030 (0.06)	0.0036 (0.08)	0.0262 (0.55)	0.0738 (1.36)
<i>TENURE</i>	-0.0017 (-0.43)	-0.0020 (-0.19)	-0.0010 (-0.11)	0.0009 (0.11)	0.0038 (0.53)	0.0036 (0.58)	-0.0007 (-0.12)	0.0012 (0.21)	0.0033 (0.53)	0.0053 (0.65)
Pseudo R ²	0.6478	0.4249	0.4702	0.5121	0.5458	0.5811	0.6189	0.6509	0.6770	0.6983

Table A2. Wald tests of coefficient equality across quantiles

This table reports F-statistics from Wald tests on simultaneous quantile regression coefficients (Table A1) with 1,000 bootstrap replications: (1) General non-linearity test (all quantiles); (2) Critical tests: lower vs. upper quantiles; (3) Tests against the median. See Appendix 1 for detailed variable definitions. ***, **, and * denote significance at 1%, 5%, and 10%, respectively.

Test Wald		TA = BTD		TA = DTAX	
		<i>BIG4</i>	<i>TENURE</i>	<i>BIG4</i>	<i>TENURE</i>
		F value	F value	F value	F value
Non-listed companies		Observations: 24,776		Observations: 16,747	
1	Q(0.10)=Q(0.20)=...=Q(0.90)	(24.12)***	(1.47)	(13.14)***	(2.09)**
2	Q(0.90) = Q(0.10)	(94.46)***	(1.95)	(70.55)***	(13.69)***
2	Q(0.80) = Q(0.20)	(28.56)***	(2.98)*	(27.73)***	(5.60)**
2	Q(0.70) = Q(0.30)	(6.37)**	(1.11)	(38.12)***	(6.89)***
3	Q(0.10) = Q(0.50)	(53.92)***	(0.52)	(24.35)***	(4.39)**
3	Q(0.20) = Q(0.50)	(0.17)	(0.03)	(6.07)**	(2.45)
3	Q(0.80) = Q(0.50)	(37.86)***	(3.45)*	(26.81)***	(4.01)**
3	Q(0.90) = Q(0.50)	(55.74)***	(1.53)	(53.27)***	(10.85)***
Listed Companies		Observations: 124		Observations: 107	
1	Q(0.10)=Q(0.20)=...=Q(0.90)	(0.10)	(0.70)	(0.72)	(0.34)
2	Q(0.90) = Q(0.10)	(0.00)	(3.11)*	(1.80)	(0.36)
2	Q(0.80) = Q(0.20)	(0.00)	(1.80)	(0.00)	(0.21)
2	Q(0.70) = Q(0.30)	(0.01)	(0.54)	(0.37)	(0.00)
3	Q(0.10) = Q(0.50)	(0.01)	(1.00)	(0.49)	(0.30)
3	Q(0.20) = Q(0.50)	(0.15)	(0.37)	(0.06)	(0.34)
3	Q(0.80) = Q(0.50)	(0.04)	(1.21)	(0.06)	(0.00)
3	Q(0.90) = Q(0.50)	(0.01)	(1.69)	(1.08)	(0.05)

By contrast, in the lower quantiles of the BTD and DTAX distributions, the association is reversed. BIG4 is negatively associated with both measures, indicating lower levels of TA among BIG4-audited firms. These findings are in line with those reported by Richardson et al. (2013) and Kanagaretnam et al. (2016). Firms located in this part of the distribution may be particularly risk-averse in fiscal matters, which would be consistent with the negative association observed for BIG4. This evidence complements the MCASHE TR results, where a similar compliance-oriented association was observed among fiscally conservative firms—an association not detected with BTG but identified with both BTD and DTAX.

In short, for non-listed firms, the association between BIG4 and TA exhibits a heterogeneous pattern across the conditional distributions of BTD and DTAX:

- In the lower quantiles (0.1–0.5 for BTD and 0.1–0.4 for DTAX), BIG4 is negatively and significantly associated with both measures, a pattern consistent with lower levels of TA among BIG4-audited firms.
- In the middle quantiles, the association becomes insignificant (0.6–0.7 for BTD and 0.5 for DTAX), representing a transition zone.
- In the upper quantiles (0.8–0.9 for BTD and 0.6–0.9 for DTAX), BIG4 is positively and significantly associated with both measures, a pattern consistent with higher levels of TA among BIG4-audited firms.

Moreover, the Wald tests confirm heterogeneity across quantiles, reveal markedly different extremes, and show that the extremes are also statistically different from the median. Therefore, these results corroborate our initial findings and reinforce support for H1 among non-listed firms, showing that the association between BIG4 and TA varies across the distribution. A positive and significant association was also observed for BTG from quantile 0.4 onwards, although it should be recalled that this measure is computed in absolute terms and therefore does not retain the direction of the book–tax difference. Nevertheless, the broadly consistent results obtained using the signed BTG and DTAX measures lend support to interpreting the positive BIG4–BTG association as being consistent with higher levels of TA. Taken together, the consistency across the three book–tax-difference-based measures reinforces the robustness of the documented distributional heterogeneity.

For listed firms (Table A1, Panels C–D), the estimates using BTG and DTAX show that BIG4 is not significantly associated with either measure at any point of the distribution. Moreover, the Wald tests do not reject coefficient equality across quantiles, providing no evidence of distributional heterogeneity. These findings reinforce the lack of support for H1 in this subsample and confirm the substantial difference between listed and non-listed firms, thereby providing further support for H2. Taken together with the main results, the evidence indicates that, among listed firms, BIG4 is not significantly associated with the book–tax-gap-based measures but is positively associated with tax payments across most of the MCASHETR distribution. This pattern is consistent with the possibility that the market scrutiny faced by listed firms strengthens BIG4 auditors’ reputational concerns in relation to the most visible dimension of tax behaviour.

For TENURE in non-listed firms, the estimates using DTAX broadly corroborate the heterogeneous pattern obtained using BTG. In the lower quantiles (0.1–0.4), TENURE is positive and significant; in the middle quantiles, it is not statistically significant; and in the highest quantile (0.9), it is negative and significant. This pattern is broadly consistent with the BTG results, where TENURE is positive and significant at quantile 0.1, insignificant across the middle of the distribution, and negative and significant at quantiles 0.8–0.9.

Moreover, the Wald tests for DTAX confirm heterogeneity across quantiles and reveal markedly different extremes. BTG provides partial support for this pattern: the coefficients at quantiles 0.2–0.4 are positive and significant, indicating that greater auditor experience is associated with higher TA among non-listed firms located in the lower part of the BTG distribution. However, although the upper-quantile BTG coefficients are negative, they are not statistically significant; thus, the sign reversal observed for BTG and DTAX is not statistically confirmed using BTG.

For listed firms, the estimates using BTG and DTAX show that TENURE is not significantly associated with TA. These results confirm our initial findings: auditor tenure shows no significant association with book–tax differences among listed firms. As previously explained, these firms are subject to market scrutiny, and their tax strategies are shaped by multiple stakeholders rather than primarily by auditor-specific knowledge.

In short, this sensitivity analysis reinforces the differences in the association between auditor tenure and TA across listed and non-listed firms. For non-listed firms, DTAX broadly corroborates the heterogeneous pattern documented using BTG, whereas BTG provides partial support by confirming the positive association in the lower part of the distribution. By contrast, TENURE shows no significant association with TA among listed firms.

Revised quantile regression for listed firms

Table A3. Estimations for Listed Companies using Three Quantiles

This table reports the results of model (1) estimated using PCSE and quantile regression at three points of the conditional TA distribution (0.25, 0.50, and 0.75). It also reports F-statistics from Wald tests. See Appendix 1 for detailed variable definitions. ***, **, and * denote significance at 1%, 5%, and 10%, respectively.

		TA = MCASHETR			TA = BTG		
		Observations: 109			Observations: 133		
		<i>BIG4</i>	<i>TENURE</i>		<i>BIG4</i>	<i>TENURE</i>	
		Coef. (t-stat)	Coef. (t-stat)	Pseudo R ²	Coef. (t-stat)	Coef. (t-stat)	Pseudo R ²
PCSE		0.402 (2.33)**	-0.098 (-2.60)***	0.460	0.093 (0.16)	-0.068 (-0.76)	0.844
0.25		0.562 (2.36)**	-0.052 (-1.37)	0.374	0.377 (0.33)	-0.043 (-0.24)	0.719
0.50		0.560 (2.11)**	-0.017 (-0.40)	0.442	0.854 (1.03)	-0.007 (-0.05)	0.734
0.75		0.570 (1.95)**	0.010 (0.12)	0.456	0.632 (1.18)	0.024 (0.21)	0.740
Test Wald		F value	F value		F value	F value	
Q(0.25)=Q(0.50)=Q(0.75)		(0.00)	(0.41)		(0.15)	(0.07)	
Q(0.25) = Q(0.75)		(0.00)	(0.63)		(0.05)	(0.13)	
Q(0.25) = Q(0.50)		(0.00)	(0.65)		(0.23)	(0.05)	
Q(0.50) = Q(0.75)		(0.00)	(0.17)		(0.10)	(0.06)	

		TA = BTD			TA = DTAX		
		Observations: 124			Observations: 107		
		<i>BIG4</i>	<i>TENURE</i>		<i>BIG4</i>	<i>TENURE</i>	
		Coef. (t-stat)	Coef. (t-stat)	Pseudo R ²	Coef. (t-stat)	Coef. (t-stat)	Pseudo R ²
PCSE		0.006 (0.31)	-0.005 (-1.13)	0.000	-0.010 (-0.45)	-0.002 (-0.43)	0.6478
0.25		0.0257 (0.44)	-0.009 (-1.13)	0.4225	0.026 (0.47)	-0.002 (-0.24)	0.4945
0.50		0.009 (0.14)	-0.005 (-0.66)	0.5059	0.062 (0.34)	0.004 (0.58)	0.5811
0.75		0.019 (0.24)	0.001 (0.08)	0.5797	0.022 (0.48)	0.003 (0.42)	0.6637
Test Wald		F value	F value		F value	F value	
Q(0.25)=Q(0.50)=Q(0.75)		(0.07)	(0.50)		(0.03)	(0.33)	
Q(0.25) = Q(0.75)		(0.01)	(0.99)		(0.00)	(0.31)	
Q(0.25) = Q(0.50)		(0.14)	(0.33)		(0.05)	(0.64)	
Q(0.50) = Q(0.75)		(0.02)	(0.53)		(0.02)	(0.04)	

To address the potential limitations associated with the small sample of listed firms, we replicate the analysis using three quantiles (0.25, 0.50, and 0.75). The nine-quantile specification provides a more detailed view of the distribution and is particularly useful for detecting non-monotonic patterns. However, in small samples, estimates for central quantiles may be more precise than those obtained at the tails of the distribution. The three-quantile specification therefore provides an additional robustness check focused on the central part of the distribution. More generally, the selection of quantiles involves balancing distributional

detail and estimation precision (Koenker & Hallock, 2001). The results of these estimations, together with the Wald tests for equality across quantiles, are presented in Table A3.

For MCASHETR, the results reinforce the homogeneous pattern observed for BIG4: the three coefficients are positive and statistically significant, while the Wald tests do not reject equality across quantiles. Therefore, the absence of heterogeneity is also observed under the three-quantile specification. In listed firms, BIG4 is consistently associated with higher MCASHETR—and therefore lower TA—across the three quantiles considered.

Using BTG, BTD, and DTAX, BIG4 is not significantly associated with any of the three book–tax-difference-based measures. For BTG, this differs from the nine-quantile results, where the coefficients were significant only at quantiles 0.8 and 0.9. However, it is consistent with the results obtained using BTD and DTAX, for which none of the BIG4 coefficients are statistically significant. Overall, the three-quantile results show no significant association between BIG4 and book–tax differences among listed firms.

For TENURE, regardless of the measure used to capture TA, no coefficient is statistically significant across the three quantiles. The Wald tests likewise do not reject equality across quantiles. Therefore, the evidence does not support H1 for listed firms. This homogeneous pattern, in contrast to the heterogeneous evidence observed among non-listed firms, reinforces H2.

Addressing Endogeneity Concerns

After conducting the main regressions and robustness checks, we obtain broadly consistent evidence regarding the associations between BIG4 and TA and between TENURE and TA, with distinct patterns for listed and non-listed firms. However, as in Armstrong et al. (2015), we cannot rule out the possibility that our estimates are affected by reverse causality or correlated omitted variables.

We considered several approaches to address these concerns. First, we explored the use of instrumental-variable quantile regression but were unable to identify an instrument that credibly satisfies both the relevance and exclusion restrictions. Second, we considered market-based instruments, such as local audit market concentration and geographic proximity to BIG4 offices. However, the data required to construct these instruments—particularly information on auditor office locations and local audit market structure—are not available for Spain. Third, we examined the possibility of exploiting the two audit reforms enacted during the study period, in 2011 (TRLAC11) and 2015 (LAC15), as regulatory shocks. However, both reforms affected all firms simultaneously and are already included as control variables in our main specifications. Finally, we considered using lagged variables as instruments. However, lagged instruments may be weak and do not necessarily satisfy the exclusion restriction when the underlying determinants are persistent over time, as may be the case with commonly used firm characteristics such as size, leverage, and profitability.

Regarding possible reverse causality, firms may select a particular auditor or maintain longer auditor tenure partly because of considerations related to their tax-planning activities. However, as in Armstrong et al. (2015), we consider tax planning to be one of many factors involved in auditor selection or retention, rather than necessarily being a primary determinant. Moreover, auditor choice tends to remain stable over long periods and is driven by multiple factors beyond

tax considerations (Francis & Wang, 2008). Especially in the case of non-listed firms, the heterogeneous associations documented for BIG4–TA and TENURE–TA across the TA distribution are less consistent with a simple explanation based on auditor selection or retention. Nevertheless, these results do not allow reverse causality to be fully ruled out.

In conclusion, while endogeneity cannot be fully eliminated, as is common in this field of research, our regression models include extensive controls for various firm characteristics and for territories with special tax regimes (FORAL and CANARIAS). In addition, we control for the two audit reforms that took place during the analysis period.

Considering all of the above, our results should be interpreted with due caution and, as Armstrong et al. (2015) note, we acknowledge that our estimates are descriptive and do not necessarily imply causality. However, quantile regression allows us to uncover patterns that would remain hidden under mean-based approaches, providing valuable insights into firm behaviour at different levels of tax aggressiveness.