Online Appendix: Audit Centralization and Audit Quality

Table OA1 presents the dynamic effects of the centralization reform on suspicious funds uncovered by city audit offices over the window of [-3, +4] where the omitted category is -1, i.e., one year before the reform. We see a distinct and sustained increase that coincides with $t = 0$ (i.e., 2016). We also note that the coefficients for pre-reform periods are statistically indistinguishable from zero, so that the parallel trend assumption required for a standard difference-in-difference test very likely holds in our setting.

Table OA2 presents the results of the centralization reform on suspicious funds uncovered by city audit offices based on the sample excluding cities in Guizhou and Yunnan. These results are similar to our main results in Table 2.

Table OA3 presents the results of the centralization reform on suspicious funds uncovered by city audit offices based on the sample including *only* cities in Guangdong as the treated group. These results are similar to our main results in Table 2.

Table OA4a compares the pre-reform (2013) city statistics for propensity score matched (PSM) sample, which is got by one-by-one nearest neighbor matching on pilot status with controls in 2013. Table OA4b presents the results of the centralization reform on suspicious funds uncovered by city audit offices based on the PSM sample. These results are similar to our main results in Table 2.

Table OA1: The Dynamic Effect on $log(Susp\,Funds)$

	(1)
Dependent Variable	$log(Susp\,Funds)$
$Pilot * Year2013$	-0.081
	(0.334)
$Pilot * Year2014$	-0.163
	(0.223)
$Pilot * Year2016$	0.484***
	(0.164)
$Pilot * Year2017$	0.460**
	(0.176)
$Pilot * Year2018$	0.276
	(0.167)
$Pilot * Year2019$	0.213
	(0.152)
$Pilot * Year2020$	0.331**
	(0.156)
Year FEs	Yes
City FEs	Yes
Controls2013*Year dummy	Yes
Observations	2650
R-Squared	0.677
Mean of Dep. Var.	12.762

Notes: Standard errors in parentheses are clustered by province. The sample covers the period from 2013 to 2020. The dependent variable in column (1) is $log(Susp\,Funds)$, which denotes the natural logarithm of suspicious fiscal funds uncovered by a city audit office. $Pilot * Year2013$ is an indicator variable denoting the third year before the centralization reform was implemented in the pilot provinces. $Pilot*Year2014$ is an indicator variable denoting the second year before the centralization reform was implemented in the pilot provinces. $Pilot * Year2016$ is an indicator variable denoting the year of the centralization reform was implemented in the pilot provinces. $Pilot * Year2017$ is an indicator variable denoting the first year after the centralization reform was implemented in the pilot provinces. $Pilot * Year2018$ is an indicator variable denoting the second year after the centralization reform was implemented in the pilot provinces. $Pilot*Year2019$ is an indicator variable denoting the third year after the centralization reform was implemented in the pilot provinces. $Pilot * Year2020$ is an indicator variable denoting the fourth year after the centralization reform was implemented in the pilot provinces. Controls in column (1) include Year indicator variables interacted with $log(GDP\,PC)$, $Industrial\,Ratio$, $log(Population)$, $log(Gov\,Rev)$, $Gov\,Balance$, FDI/GDP, and $Avg\,Education$ in 2013. The coefficients and standard errors of the control variables are suppressed to conserve space. See Appendix Table A1 for detailed definitions of the control variables.
Significance: * significant at 10%; ** significant at 5%; *** significant at 1%.

Table OA2: Excluding Cities in Guizhou and Yunnan

	(1)	(2)	(3)	(4)	(5)
Dependent Variable			$log(Susp\,Funds)$		
$Pilot * Post2016$	0.319*	0.429**	0.375**	0.375**	0.186
	(0.168)	(0.166)	(0.156)	(0.147)	(0.164)
Year FEs	Yes	Yes	Yes	Yes	Yes
City FEs	Yes	Yes	Yes	Yes	Yes
Controls2013*Time trend		Yes			
Controls2013*Post2016			Yes		
Controls2013*Year dummy				Yes	
Controls					Yes
Observations	2450	2450	2450	2450	2143
R-Squared	0.664	0.670	0.668	0.682	0.681
Mean of Dep. Var.	12.773	12.773	12.773	12.773	12.739

Notes: Standard errors in parentheses are clustered by province. The sample covers the period from 2013 to 2020 and excludes cities in Guizhou and Yunnan. The dependent variable in columns (1)-(5) is $log(Susp\,Funds)$, which denotes the natural logarithm of suspicious fiscal funds uncovered by a city audit office. $Pilot$ is an indicator variable denoting pilot provinces and $Post2016$ denotes years 2016 and later after the centralization reform was implemented. Controls in column (2) include Time trend interacted with $log(GDP\,PC)$, $Industrial\,Ratio$, $log(Population)$, $log(Gov\,Rev)$, $Gov\,Balance$, FDI/GDP, and $Avg\,Education$ in 2013. Controls in column (3) include $Post2016$ interacted with $log(GDP\,PC)$, $Industrial\,Ratio$, $log(Population)$, $log(Gov\,Rev)$, $Gov\,Balance$, FDI/GDP, and $Avg\,Education$ in 2013. Controls in column (4) include Year indicator variables interacted with $log(GDP\,PC)$, $Industrial\,Ratio$, $log(Population)$, $log(Gov\,Rev)$, $Gov\,Balance$, FDI/GDP, and $Avg\,Education$ in 2013. Controls in column (5) include $log(GDP\,PC)$, $Industrial\,Ratio$, $log(Population)$, $log(Gov\,Rev)$, $Gov\,Balance$, FDI/GDP, and $Avg\,Education$. The coefficients and standard errors of the control variables are suppressed to conserve space. See Appendix Table A1 for detailed definitions of the control variables.

Significance: * significant at 10%; ** significant at 5%; *** significant at 1%.

Table OA3: Including Only Cities in Guangdong as the Treated Group

	(1)	(2)	(3)	(4)	(5)
Dependent Variable			$log(Susp\,Funds)$		
$Pilot * Post2016$	0.674***	0.745***	0.696***	0.679***	0.548***
	(0.087)	(0.115)	(0.114)	(0.115)	(0.117)
Year FEs	Yes	Yes	Yes	Yes	Yes
City FEs	Yes	Yes	Yes	Yes	Yes
Controls2013*Time trend		Yes			
Controls2013*Post2016			Yes		
Controls2013*Year dummy				Yes	
Controls					Yes
Observations	2124	2124	2124	2124	1857
R-Squared	0.625	0.634	0.631	0.647	0.647
Mean of Dep. Var.	12.577	12.577	12.577	12.577	12.542

Notes: Standard errors in parentheses are clustered by province. The sample covers the period from 2013 to 2020 and excludes cities in Jiangsu, Zhejiang, Shandong, Guizhou, and Yunnan. The dependent variable in columns (1)-(5) is $log(Susp\,Funds)$, which denotes the natural logarithm of suspicious fiscal funds uncovered by a city audit office. *Pilot* is an indicator variable denoting pilot provinces and $Post2016$ denotes years 2016 and later after the centralization reform was implemented. Controls in column (2) include Time trend interacted with $log(GDP\,PC)$, *Industrial Ratio*, $log(Population)$, $log(Gov\,Rev)$, *Gov Balance*, FDI/GDP, and *Avg Education* in 2013. Controls in column (3) include $Post2016$ interacted with $log(GDP\,PC)$, *Industrial Ratio*, $log(Population)$, $log(Gov\,Rev)$, *Gov Balance*, FDI/GDP, and *Avg Education* in 2013. Controls in column (4) include Year indicator variables interacted with $log(GDP\,PC)$, *Industrial Ratio*, $log(Population)$, $log(Gov\,Rev)$, *Gov Balance*, FDI/GDP, and *Avg Education* in 2013. Controls in column (5) include $log(GDP\,PC)$, *Industrial Ratio*, $log(Population)$, $log(Gov\,Rev)$, *Gov Balance*, FDI/GDP, and *Avg Education*. The coefficients and standard errors of the control variables are suppressed to conserve space. See Appendix Table A1 for detailed definitions of the control variables.
Significance: * significant at 10%; ** significant at 5%; *** significant at 1%.

Table OA4a: Comparison of Baseline (2013) City Statistics for Propensity Score Matched Sample

	Pilot		Non Pilot		Pilot vs. Non	
Variables	Mean	StdDev	Mean	StdDev	Difference	t-statistic
$log(GDP\,PC)$	10.614	0.841	10.600	0.569	0.014	0.067
$Industrial\ Ratio$	0.478	0.079	0.460	0.109	0.018	0.651
$log(Population)$	5.888	0.663	5.825	0.811	0.063	0.310
$log(Gov\,Rev)$	14.024	0.994	13.951	1.041	0.073	0.233
$Gov\,Balance$	2.322	1.428	2.583	2.000	-0.261	-0.437
FDI/GDP	0.017	0.017	0.016	0.016	0.001	0.297
$Avg\,Education$	9.887	0.444	9.734	1.115	0.153	0.487

Notes: Pilot indicates the group of each reform province. Non Pilot indicates the group of each non-reform province. See Appendix Table A1 for detailed definitions of the variables. T-statistics are calculated based on a regression of each outcome, clustering at the province level.

Table OA4b: Estimates based on Propensity Score Matched Sample

	(1)	(2)	(3)	(4)	(5)
Dependent Variable			$log(Susp\,Funds)$		
$Pilot * Post2016$	0.379*	0.417**	0.428**	0.443***	0.301*
	(0.195)	(0.170)	(0.156)	(0.151)	(0.157)
Year FEs	Yes	Yes	Yes	Yes	Yes
City FEs	Yes	Yes	Yes	Yes	Yes
Controls2013*Time trend		Yes			
Controls2013*Post2016			Yes		
Controls2013*Year dummy				Yes	
Controls					Yes
Observations	1016	1016	1016	1016	889
R-Squared	0.610	0.629	0.631	0.659	0.635
Mean of Dep. Var.	13.206	13.206	13.206	13.206	13.172

Notes: Standard errors in parentheses are clustered by province. The sample covers the period from 2013 to 2020 and is propensity score matched on $Pilot$ by one-by-one nearest neighbor matching with controls in 2013. The dependent variable in columns (1)-(5) is $log(Susp\,Funds)$, which denotes the natural logarithm of suspicious fiscal funds uncovered by a city audit office. $Pilot$ is an indicator variable denoting pilot provinces and $Post2016$ denotes years 2016 and later after the centralization reform was implemented. Controls in column (2) include Time trend interacted with $log(GDP\,PC)$, $Industrial\,Ratio$, $log(Population)$, $log(Gov\,Rev)$, $Gov\,Balance$, FDI/GDP, and $Avg\,Education$ in 2013. Controls in column (3) include $Post2016$ interacted with $log(GDP\,PC)$, $Industrial\,Ratio$, $log(Population)$, $log(Gov\,Rev)$, $Gov\,Balance$, FDI/GDP, and $Avg\,Education$ in 2013. Controls in column (4) include Year indicator variables interacted with $log(GDP\,PC)$, $Industrial\,Ratio$, $log(Population)$, $log(Gov\,Rev)$, $Gov\,Balance$, FDI/GDP, and $Avg\,Education$ in 2013. Controls in column (5) include $log(GDP\,PC)$, $Industrial\,Ratio$, $log(Population)$, $log(Gov\,Rev)$, $Gov\,Balance$, FDI/GDP, and $Avg\,Education$. The coefficients and standard errors of the control variables are suppressed to conserve space. See Appendix Table A1 for detailed definitions of the control variables.
Significance: * significant at 10%; ** significant at 5%; *** significant at 1%.