

Supplementary material

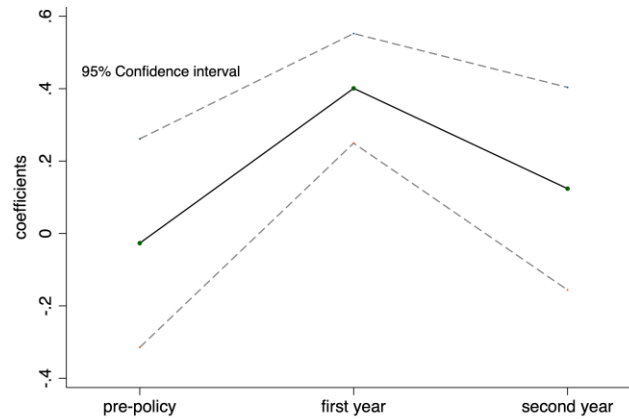
Supplementary file Table A0. Summary statistics of control variables in natural scale (N=1680)

	Obs	Mean	Std.	dev.	Min
Panel A: Full sample					
<i>wage of all workers</i>	1,677	2284069	2872163	174391	2.60e+07
<i>per capita GDP</i>	1,680	46675.61	30192.88	6457	290477
<i>education expenditure</i>	1,680	525758.5	378507.2	20048	3219820
<i>population</i>	1,680	424.5666	253.5467	19.5	1399
Panel B: 2011-2014					
<i>wage of all workers</i>	1,115	1996167	2474037	174391	2.30e+07
<i>per capita GDP</i>	1,118	44350.57	30507.6	6457	290477
<i>education expenditure</i>	1,118	474200.5	326802.7	20048	2539473
<i>population</i>	1,118	423.2512	251.2843	19.5	1238.5
Panel C: 2015-2016					
<i>wage of all workers</i>	562	2855264	3463958	213157	2.60e+07
<i>per capita GDP</i>	562	51300.87	29034.36	10987	215488
<i>education expenditure</i>	562	628324.2	447546.2	35925	3219820
<i>population</i>	562	427.1834	258.1945	20.25	1399

Notes: The units of the variables are as follows: Total wages: 10,000 yuan; GDP per capita: yuan; Education expenditure: 10,000 yuan; Total population: 10,000 people.

The observed policy effect of GST increase might not be attributable solely to the policy itself but might be also influenced by other observed and unobserved factors. To address this potential threat we additionally employ several different methods, and the results demonstrate that our conclusions are still robust.

First, when employing the DID method, the outcome variable (SIVST) among cities should be indifference before the implementation of the policy. The parallel trend test graph shows that before the policy implementation, cities with large per capita smuggling scales does not exhibit significant larger smuggling scales compared to cities with smaller per capita smuggling scales. This means that after the policy implementation, the significant difference of smuggling between cities stems from the GST increase.



Supplementary file Figure A1. The difference of SIVST among cites before and after the policy implementation

Second, in a linear panel event-study design, unobserved confounds may be related both to the outcome and to the policy variable of interest. To address this problem, a placebo test and a Difference-in-Difference-in-Difference (DDD) method are used. By using fake different cut-off points of 2013 for the policy implementation period, we redefine the key explanatory variable as *Exposure*dummy_{fake}* and re-estimate model. Results are shown in column (1) of Table A1, and the coefficient of the interaction term is no longer significant. This finding ensures that the observed effects are not due to other policy or confounding factors. By classifying sample into two groups based on cigarette sales of cities in 2014, we define a dummy variable *Sale* that equals 1 if cities with higher cigarette sales, otherwise, 0. We introduce the triple interaction of *Sale* and original key explanatory variable into regression model to eliminate the possible pre-event trend caused by unobserved factors. Result in column (2) shows that coefficient of the triple interaction is significantly at 1%, which further validates our conclusions.

Supplementary file Table A1. The effect of GST increase on SIVST among cities after delete observations of each years, respectively

<i>SIVST</i>	(1)	(2)
<i>Exposure*dummy_{fake}</i>	43.66 (1.09)	
<i>Exposure*dummy₂₀₁₅*Sale</i>		0.277*** (3.02)
<i>Log wage of all workers</i>	2.123 (0.79)	2.092 (0.90)
<i>Log per capita GDP</i>	1.905 (0.58)	1.664 (0.53)
<i>Log education expenditure</i>	5.050* (1.73)	4.759* (1.71)
<i>Log population</i>	2.189 (0.29)	3.062 (0.42)
<i>constant</i>	-113.3* (-1.80)	-111.8* (-1.83)
CITY_FE	YES	YES
YEAR_FE	YES	YES
R ²	0.03	0.03
N	1622	1642

Notes: the t statistics are in parentheses. Standard error clustered at the prefecture level. *, **, *** statistically distinct from 0 at the 10%, 5%, 1% significance level, respectively. CITY_FE and YEAR_FE indicate city-fixed effects and year-fixed effects, respectively. R2 is a statistic that reflects the model's goodness of fit as the ratio of the regression sum of squares to the total sum of squares. N denotes the sample size for the regression analysis.

Third, employment situation matters to economic crimes, thus it is necessary to consider them as control variables. Specifically, we incorporate two variables *Employment* (ratio of employment

to total population) and *Unemployment* (ratio of unemployment to total population) reflecting the employment situation of cities as control variables in the model. Results are reported in Table A2. Columns (1) to (3) show positively significant coefficients of the interaction term. That is, our conclusions are unchanged under the controlling of employment situation.

Supplementary file Table A2. The effect of GST increase on SIVST after controlling employment situation and improving the definition of key explanatory variable

<i>SIVST</i>	(1)	(2)	(3)	(4)	(5)
<i>Exposure*dummy₂₀₁₅</i>	0.252*** (4.15)	0.267*** (4.50)	0.266*** (4.47)		
<i>Exposure*dummy_{new}</i>				39.48* (1.85)	59.18*** (2.74)
<i>Log wage of all workers</i>	3.753 (1.33)	2.260 (0.96)	3.993 (1.39)	1.698 (0.70)	4.668 (1.60)
<i>Log per capita GDP</i>	1.631 (0.52)	0.819 (0.26)	1.169 (0.38)	3.347 (1.04)	4.105 (1.29)
<i>Log education expenditure</i>	5.068* (1.88)	5.766** (1.97)	5.557* (1.92)	4.919* (1.75)	5.083* (1.69)
<i>Log population</i>	1.342 (0.18)	1.766 (0.23)	0.512 (0.07)	1.493 (0.20)	-3.588 (-0.49)
<i>Employment</i>	-14.04** (-2.17)		-14.24** (-2.15)		-24.26*** (-3.51)
<i>Unemployment</i>		0.00308 (0.41)	0.00354 (0.48)		0.00238 (0.31)
<i>constant</i>	-126.6** (-2.06)	-110.5* (-1.79)	-126.6** (-2.04)	-116.6* (-1.85)	-135.3** (-2.12)
CITY_FE	YES	YES	YES	YES	YES
YEAR_FE	YES	YES	YES	YES	YES
R ²	0.03	0.04	0.04	0.03	0.03

N	1642	1622	1622	1642	1622
---	------	------	------	------	------

Notes: the t statistics are in parentheses. Standard error clustered at the prefecture level. *, **, *** statistically distinct from 0 at the 10%, 5%, 1% significance level, respectively. CITY_FE and YEAR_FE indicate city-fixed effects and year-fixed effects, respectively. R2 is a statistic that reflects the model's goodness of fit as the ratio of the regression sum of squares to the total sum of squares. N denotes the sample size for the regression analysis.

Finally, considering the correlation between pre-policy SIVST and post-policy SIVST, and to demonstrate that our findings are insensitive to the construction way of key explanatory variable, we use a more robust indicator to substitute the original key explanatory variable. We improve the original key explanatory variable by substituting the cigarette sales amount of cities in 2014 for $SIVST_{2014}$ in equation (2) of manuscript for two reasons: 1) The amount of cigarette sales in cities, to some extent, is exogenous with SIVST of cities. The insignificant coefficients of the interaction term and SIVST in Table A3 show that not only SIVST is uncorrelated with cigarette sales, but also the 2015 cigarette tax policy affect little on cigarette sales. This indicates that cigarette sales are independent of smuggling activities and the increase in GST, thereby satisfying the exogeneity requirements for the variable construction of policy shock. 2) Sales indicate market demand and reflect both the scale and difficulty of profiting from smuggling. Following the tax increase, cities with higher sales offer offenders greater profit opportunities through smuggling. Therefore, sales volume directly reflects the intensity of the impact of policy. Thus, it is reasonable to use the cigarette sales of cities as an proxy variable of policy. Coefficients of the new key explanatory variable are shown in columns (4) and (5) of Table A2. The results demonstrate that our conclusions are still robust.

Supplementary file Table A3. The relationship between SIVST and cigarette sales

Sales	(1)	(2)
$Exposure*dummy_{2015}$	0.0182 (0.74)	
$SIVST$		0.00262 (0.15)

<i>Log wage of all workers</i>	5.231*** (3.36)	5.168*** (3.36)
<i>Log per capita GDP</i>	4.166** (2.07)	3.720* (1.78)
<i>Log education expenditure</i>	4.142*** (3.00)	4.330*** (3.07)
<i>Log population</i>	9.754 (1.09)	6.881 (0.83)
<i>constant</i>	-197.9*** (-3.75)	-178.1*** (-3.70)
CITY_FE	YES	YES
YEAR_FE	YES	YES
R ²	0.41	0.42
N	1673	1642

Notes: the *t* statistics are in parentheses. Standard error clustered at the prefecture level. *, **, *** statistically distinct from 0 at the 10%, 5%, 1% significance level, respectively. CITY_FE and YEAR_FE indicate city-fixed effects and year-fixed effects, respectively. R² is a statistic that reflects the model's goodness of fit as the ratio of the regression sum of squares to the total sum of squares. N denotes the sample size for the regression analysis.

Supplementary file Table A4. Heterogeneity effect of GST increase on economic crime between central and coastal cities (N=1545)

	(1)	(2)	(3)
	<i>SIVST</i>	<i>SIVST</i>	<i>SIVST</i>
<i>Exposure*dummy₂₀₁₅*coastal</i>	1.511** (2.01)	1.348* (1.77)	1.305* (1.68)
<i>Exposure*dummy₂₀₁₅</i>	0.287*** (3.86)	0.253*** (3.32)	0.246*** (3.19)
<i>Log wage of all workers</i>			1.492

			(0.60)
<i>Log per capita GDP</i>			2.288
			(0.73)
<i>Log education expenditure</i>			4.174
			(1.43)
<i>Log population</i>			-0.802
			(-0.11)
<i>constant</i>	12.80***	13.02***	-79.60
	(76.74)	(16.12)	(-1.28)
CITY_FE	YES	YES	YES
YEAR_FE	NO	YES	YES
R ²	0.03	0.04	0.04
N	1549	1549	1545

Notes: the *t* statistics are in parentheses. Standard error clustered at the prefecture level. *, **, *** statistically distinct from 0 at the 10%, 5%, 1% significance level, respectively. CITY_FE and YEAR_FE indicate city-fixed effects and year-fixed effects, respectively. R² is a statistic that reflects the model's goodness of fit as the ratio of the regression sum of squares to the total sum of squares. N denotes the sample size for the regression analysis.

Supplementary file Table A5. Heterogeneity effect of GST increase on economic crime between central and border cities (N=1371)

	(1)	(2)	(3)
	<i>SIVST</i>	<i>SIVST</i>	<i>SIVST</i>
<i>Exposure*dummy2015*border</i>	-0.0489	-0.0366	-0.0338
	(-0.60)	(-0.45)	(-0.39)
<i>Exposure*dummy2015</i>	0.287***	0.256***	0.244***
	(3.86)	(3.46)	(3.22)
<i>Log wage of all workers</i>			3.433
			(1.31)

<i>Log per capita GDP</i>			3.400
			(1.04)
<i>Log education expenditure</i>			4.696**
			(2.04)
<i>Log population</i>			1.114
			(0.20)
<i>constant</i>	11.30***	10.66***	-137.1***
	(130.33)	(14.38)	(-2.80)
CITY_FE	YES	YES	YES
YEAR_FE	NO	YES	YES
R ²	0.02	0.03	0.03
N	1375	1375	1371

Notes: the *t* statistics are in parentheses. Standard error clustered at the prefecture level. *, **, *** statistically distinct from 0 at the 10%, 5%, 1% significance level, respectively. CITY_FE and YEAR_FE indicate city-fixed effects and year-fixed effects, respectively. R² is a statistic that reflects the model's goodness of fit as the ratio of the regression sum of squares to the total sum of squares. N denotes the sample size for the regression analysis.

© 2025 Zhang X. and Zhang Z.