Electoral Impacts of a Failed Uprising: Evidence from Hong Kong's Umbrella Movement
(Online Appendix)

Appendix A: Our Estimation Strategies

The DID estimators

In this section, we explain the methods we used in the empirical analysis. As indicated in the main text, our estimation strategy is difference-in-differences (DID). For results on electoral support, we estimate the following equation:

$$Dem\_VoteShare_{ct} = \mu + \beta \mathbb{1}\{PostUmbrella\} * Distance_c + \gamma \mathbb{1}\{PostUmbrella\} + \alpha_c + \epsilon_{ct}. \quad (1)$$

The unit of analysis is constituency c and the outcome is the vote share of the pro-democracy opposition in constituency c and year t, $Dem\_VoteShare_{ct}$. The year t is either 2012 or 2016. As we have controlled the fixed effects of constituencies, α_c, there is no need to add a separate term for $Distance_c$ in the regression equation, for its effect will be absorbed by α_c. After taking the first difference on both sides of Equation (1), all the time-invariant variables are eliminated and we have

$$\Delta Dem\_VoteShare_{ct} = \gamma + \beta Distance_c + \Delta\epsilon_{ct}. \quad (2)$$

Following this equation, we can plot how the change in the vote share of the opposition varies with the distance of each constituency c to the protest sites, as in Figure 2 of the main text. For the results on turnout, we simply replace the vote share of the opposition with the turnout rate in constituency c, year t. For the placebo test, we use election results in 2008 and 2012, rather than 2012 and 2016.

When analyzing mechanisms, we use a slightly different model specification, because the ABS survey is not a panel study and we possess only the geo-locations of respondents at the district level. The model is as follows:

$$Indicator_{idt} = \mu + \beta \mathbb{1}\{PostUmbrella\} * Distance_d + \gamma \mathbb{1}\{PostUmbrella\} + \delta \mathbf{X}_{idt} + \alpha_d + \epsilon_{idt}, \quad (3)$$

where i represents individuals and d denotes districts; the outcome variable $Indicator_{idt}$ is the indicator constructed via PCA that corresponds to each of the three proposed mechanisms (perceived political efficacy, approval for democracy, and perceived economic insecurity) for individual i in district d, year t. Notice that we can only include district fixed effects α_d rather than individual fixed effects α_i, since each individual is observed only once in the sample. But in this case, we can control for covariates at individual level as they won't be absorbed by district fixed effects. Now, after taking the first difference, what we have is:

$$Indicator_{id,2016} - \overline{Indicator}_{d,2012} = \gamma + \beta Distance_d + \delta \mathbf{X}_{id,2016} - \bar{\mathbf{X}}_{d,2012} + \epsilon_{id,2016} - \bar{\epsilon}_{d,2012}, \quad (4)$$

where $\overline{Indicator}_{d,2012}$ is the 2012 average value of the indicator in district d where individual i resides; $\bar{\mathbf{X}}_{d,2012}$ and $\bar{\epsilon}_{d,2012}$ are defined similarly. In other words, we difference out district-level averages rather than individual level ones. Clearly, via this approach, we are unable to eliminate all the time-invariant confounders at the individual level. But the concern is partly alleviated by the fact that the inclusion of $\mathbf{X}_{idt}$ into the regression does not change the estimates much (see the comparison between Tables 3A, 3B in the main text and Tables A5, A6, and A7 below). To understand the average change in these indicators, we also estimate γ in Equation (4) without the $\beta Distance_d$ term. This is the specification that has no interaction in Tables 4A, 4B in the main text and Tables A5, A6, and A7 below. When plotting Figure

A5 below, we take the district-level average of Equation (4) and plot $\overline{Indicator}_{d,2016} - \overline{Indicator}_{d,2012}$ against $Distance_d$.

Finally, for the individual-level DID analysis in Table 5 of the main text, we use the model specification below:

$$Dem\_VoteShare_{idt} = \mu + \beta \mathbb{1}\{PostUmbrella\} * Distance_d + \gamma \mathbb{1}\{PostUmbrella\} + \alpha_i + \epsilon_{idt}. \quad (5)$$

Now we have individual fixed effects as the HKES survey is indeed panel data. Taking the first difference of Equation (5), we have:

$$\Delta Dem\_VoteShare_{idt} = \gamma + \beta Distance_d + \Delta\epsilon_{idt}. \quad (6)$$

Compared with Equation (2), there are two differences. First, $Dem\_VoteShare_{idt}$ only takes two values, 1 if voting for the opposition and 0 if voting for the pro-establishment camp. Second, the distance is measured at district rather than constituency level. Therefore, both the outcome and the explanatory variable are measured less accurately. But our estimates are significant even in this scenario.

Control the covariates

The covariates at constituency level that we use in the paper are obtained from the census in Hong Kong, which is taken every 10 years. Therefore, the values of these covariates remain unchanged between 2012 and 2016. We use both regression and weighting to account for the influence of these covariates. The regression method is equivalent to adding an extra term $\delta \mathbf{X}_c$ to the right-hand side of Equation (2). As presented in Table A3, we try two approaches to decide what covariates to include. The first one is based on our substantive knowledge and the second one exploits a machine learning algorithm known as LASSO (least absolute shrinkage and selection operator) that selects covariates to maximize the model's power of prediction. The two approaches lead to different sets of covariates, but our main results remain statistically significant under both of them.

The weighting method we rely on is CBPS for continuous treatment proposed by Fong et al. (2018). The basic idea is to find w_i for each unit i in the sample, such that the weighted averages of each standardized covariate X_i, treatment D_i, and their interaction are equal to zero. They show in the paper that the weighting algorithm creates a sample that is balanced in all the covariates along each value of the treatment. We consider $Distance_c$ as the treatment and re-estimate Equation (2) after weighting. The results are presented in Figure 3 of the main text. We also have two time-varying covariates at district level that are reported in Hong Kong's by-census, which is conducted every 5 years. Results with these two covariates are reported in columns 2 and 5 of Table A3.

Sensitivity test

The sensitivity test is an effective tool for checking the potential influence of unobservable confounders on the estimation result. The basic idea is to assume the existence of a confounder U that is correlated with both the treatment and the outcome. We vary these two correlation coefficients to see when the result will be driven to be zero or insignificant. We follow the approach proposed in Carnegie, Harada, and Hill (2016) as it is more flexible than other alternatives.

Results of the sensitivity test are presented in Figure A3. The left graph is for the super-seat election and the right one for the regular-seat election. The x-axis in each graph represents the correlation between the unobservable confounders and the treatment, ζ^z. Similarly, the y-axis marks the correlation between the unobservable confounders and the outcome,[1] ζ^y. ζ^z and ζ^y reflect the influence of the unobservables on the treatment and the outcome, respectively. For each pair (ζ^z, ζ^y), we can calculate the estimate for the treatment effect and its p-value. Contours on the graphs indicate the combinations of (ζ^z, ζ^y) that would lead to the same treatment effect estimate (labeled on each contour). Among all the contours, the red ones are where the estimated effect would be zero and the blue ones are where it turns insignificant.

Red plus signs and blue triangles show the correlation coefficients between each observable covariate and the treatment as well as the outcome.[2] As may be seen from the figures, the unobservables must be larger than any of the observable to render the variable of interest insignificant. In the super-seat

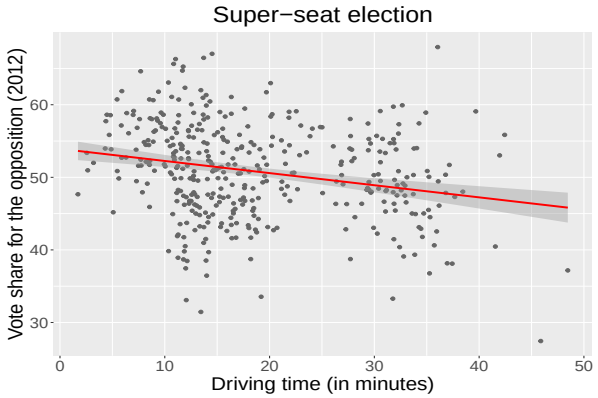
[1] Without loss of generality, we assume that the unobservables are positively correlated with the outcome.

[2] The red plus sign suggests that the covariate actually has a negative correlation with the outcome. But for simplicity the sign is flipped over. It will not affect the result as only the strength of the correlation matters, not the direction.

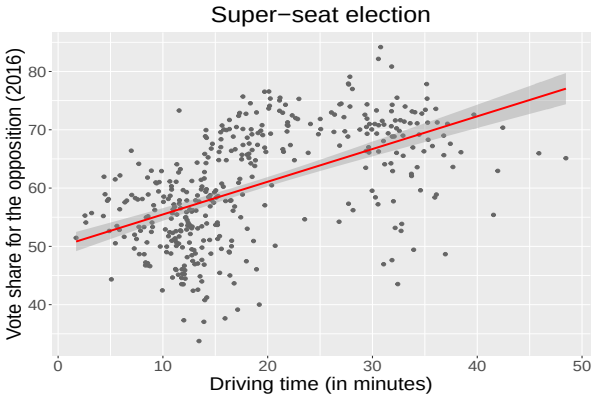
election, this happens when both ζ^z and ζ^y are larger than 0.5. As for the regular-seat election, the estimate becomes insignificant when both ζ^z and ζ^y are larger than 0.3. They suggest that our estimates would not be insignificant unless the unobservables were as predictive as the most predictive covariate, which is quite unlikely in practice.

Appendix B: Graphs

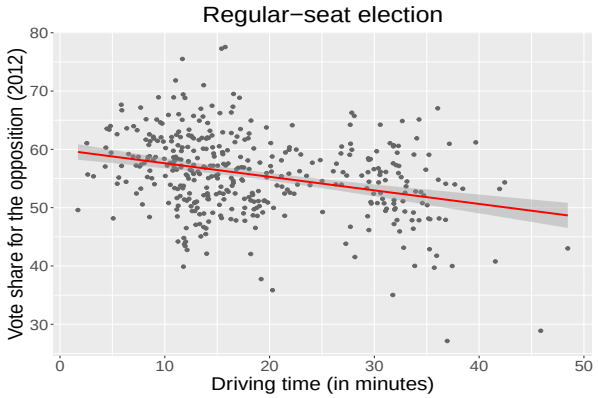
Figure A1: The relationship between the vote share of the opposition and proximity to protest sites



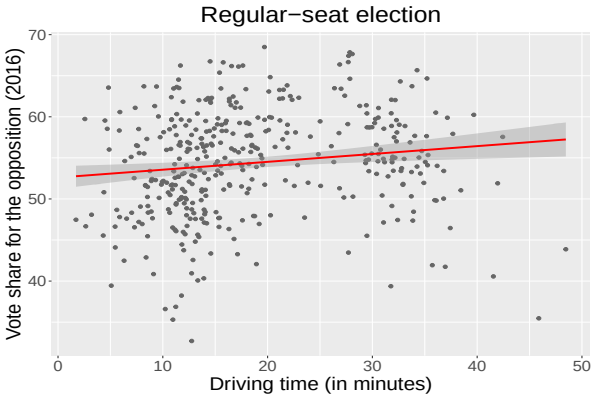
(a) Vote share for the opposition (2012, super-seat election)



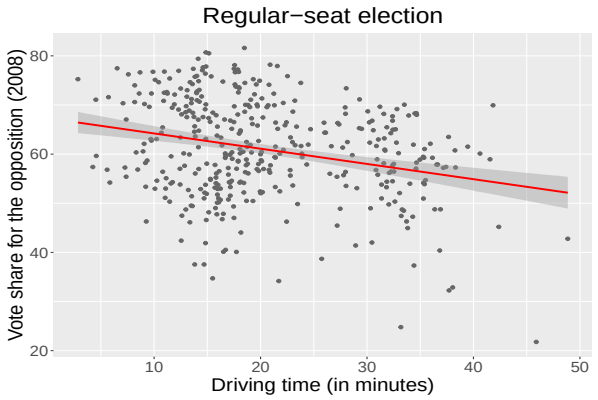
(b) Vote share for the opposition (2016, super-seat election)



(c) Vote share for the opposition (2016, regular-seat election)



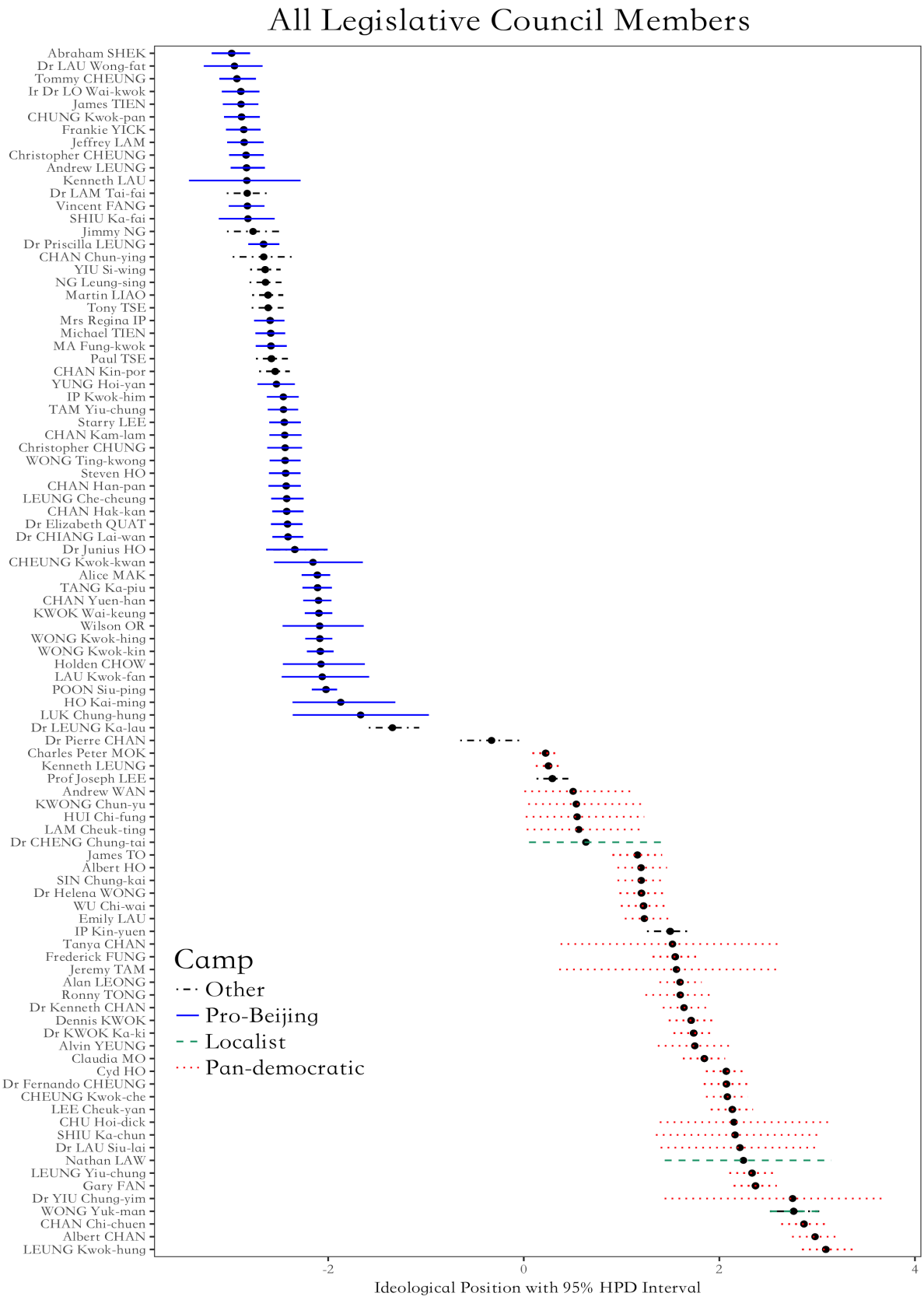
(d) Vote share for the opposition (2012, regular-seat election)



(e) Vote share for the opposition (2008, super-seat election)

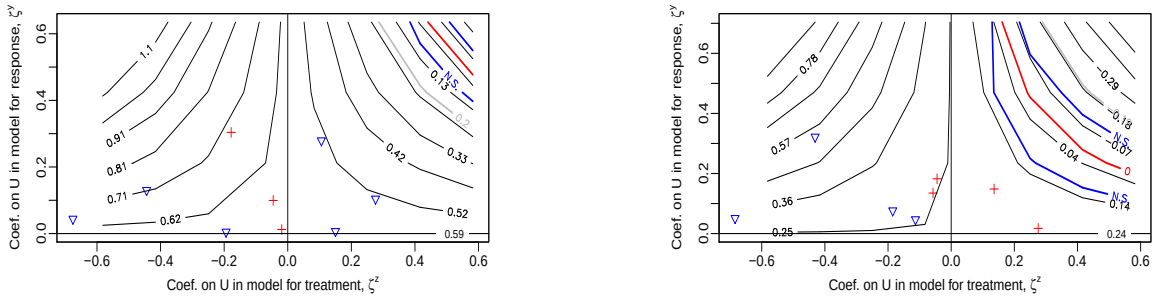
Notes: The figures show how the vote share of the opposition varies with the distance to the protest sites in the 2008, 2012 and 2016 Legislative Council elections. Each black spot in the graphs represents the vote share in the corresponding constituency. The red line is the linear fit of the spots, with the grey area indicating its 95% confidence interval. The first row is for the super-seat election and the next two are for the regular-seat election. The two columns display the relationship between the vote share and the distance before and after the Umbrella Movement, respectively. It is noteworthy that the general pattern reverses after the protest.

Figure A2: Ideal points of Hong Kong legislators in the 5th and 6th Legislative Councils



Notes: The figure displays the estimated ideal points of all Hong Kong legislators in the 5th and 6th Legislative Councils (2012-2017). Black spots mark the position of each legislator on the ideological spectrum; colored lines indicate the range of confidence interval. Different colors represent different political camps. The estimation is conducted using the R package "emIRT" on the basis of legislators' roll-call voting records.

Figure A3: Results of sensitivity test

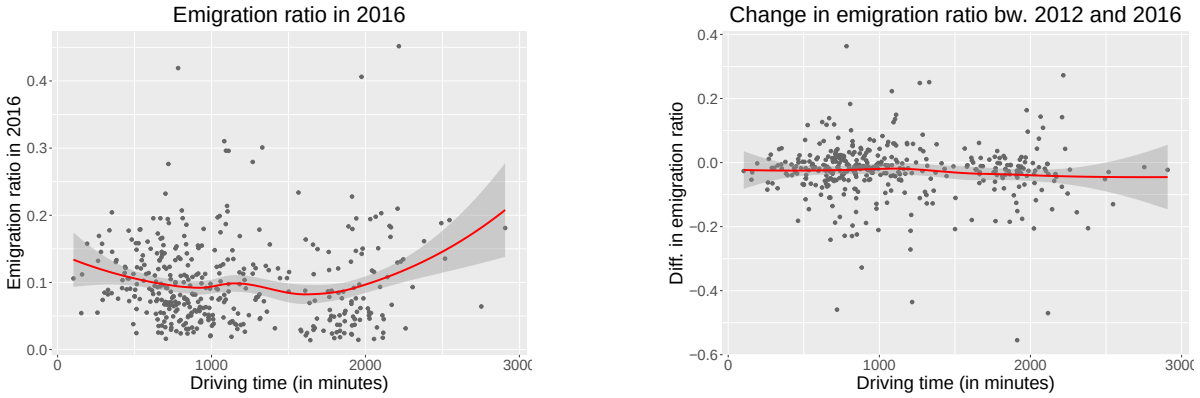


(a) Super-Seat Election

(b) Regular-Seat Election

Notes: The sensitivity tests follow the approach proposed in Carnegie, Harada, and Hill (2016). The left graph is for the super-seat election and the right one is for the regular-seat election. The x-axis and y-axis indicate the correlation of the unobservables with the treatment and the outcome, respectively. Each contour on the graphs reflects the combinations of correlation coefficients that would lead to the same treatment effect estimate (labeled on the contour). The red contours are where the estimated treatment effect becomes zero and the blue contours are where it turns insignificant. Red plus signs and blue triangles give the correlation coefficients of observable covariates with the treatment and the outcome.

Figure A4: Pattern of emigration after the protest

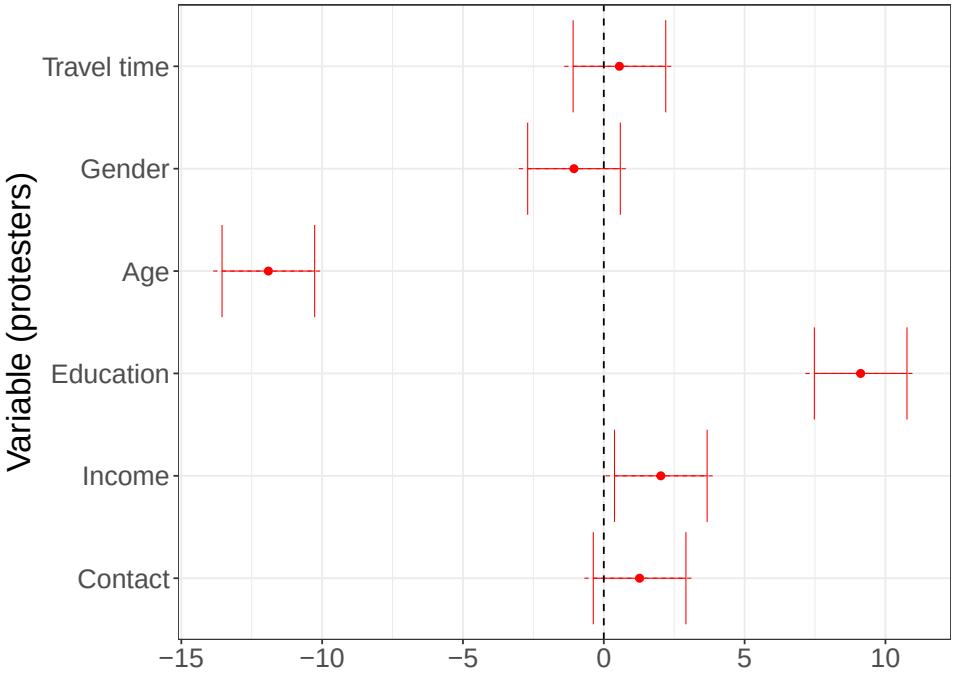


(a) Emigration ratio in 2016

(b) Change in emigration ratio between 2012 and 2016

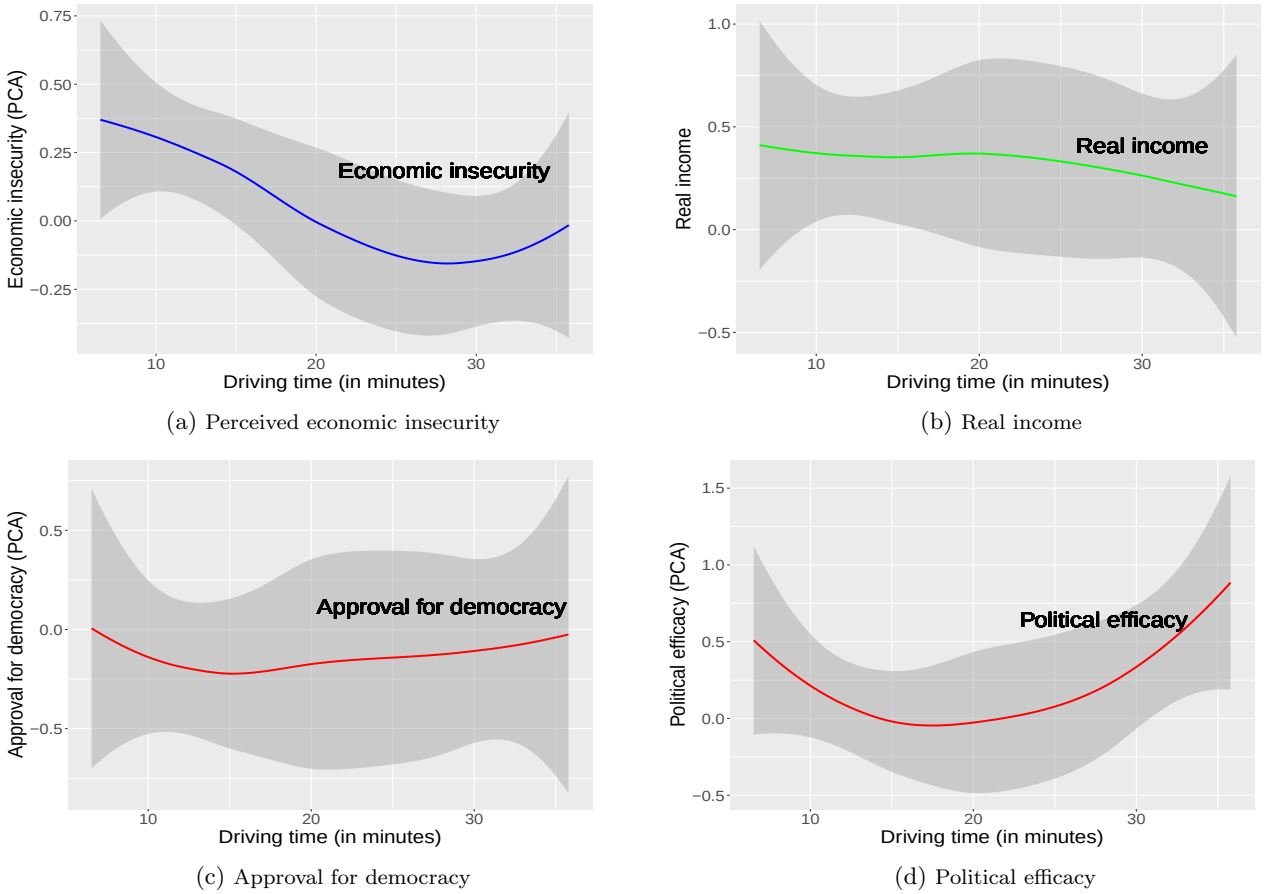
Notes: The figures show how the pattern of emigration in each constituency varies with the distance to the occupy sites. On the left the outcome is the emigration ratio in 2016, and on the right it is the change in emigration ratio between 2012 and 2016. Each black spot in the graphs represents the (change of) emigration ratio in the corresponding constituency. The red line is the linear fit of the spots, with the grey area indicating its 95% confidence interval. It is noteworthy that neither the emigration ratio in 2016 nor its change has a significant relationship with proximity to protest sites.

Figure A5: Comparison of protesters and non-protesters



Notes: This figure is based Asian Barometer Survey, Wave 4. The classification of protesters and non-protesters relies on their self-reported participation in the Umbrella Movement. Each point indicates the average difference between the two groups. The solid line represents the 90% confidence interval of the difference, and the dotted line represents the 95% confidence interval.

Figure A6: Relationship between PCA scores and proximity to protest sites



(a) Perceived economic insecurity

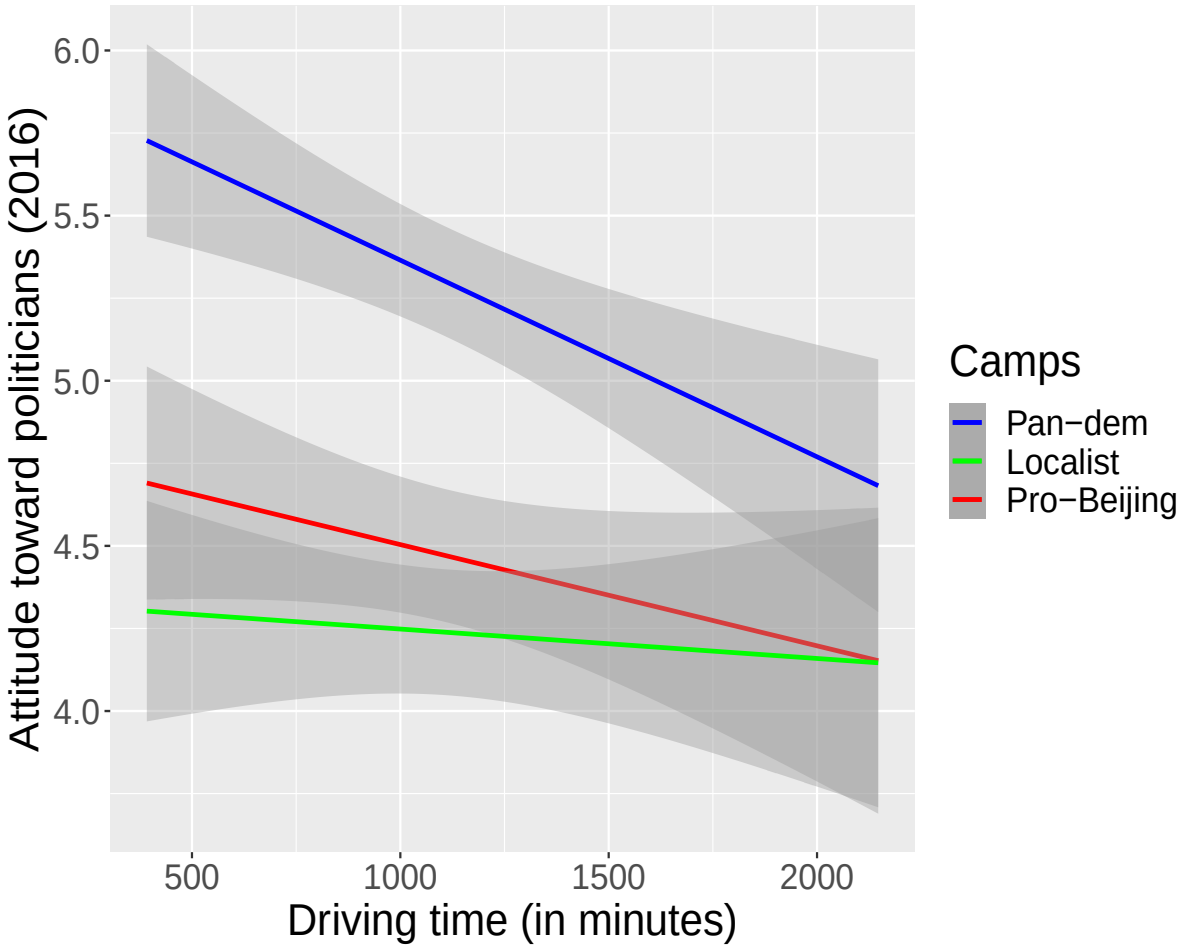
(b) Real income

(c) Approval for democracy

(d) Political efficacy

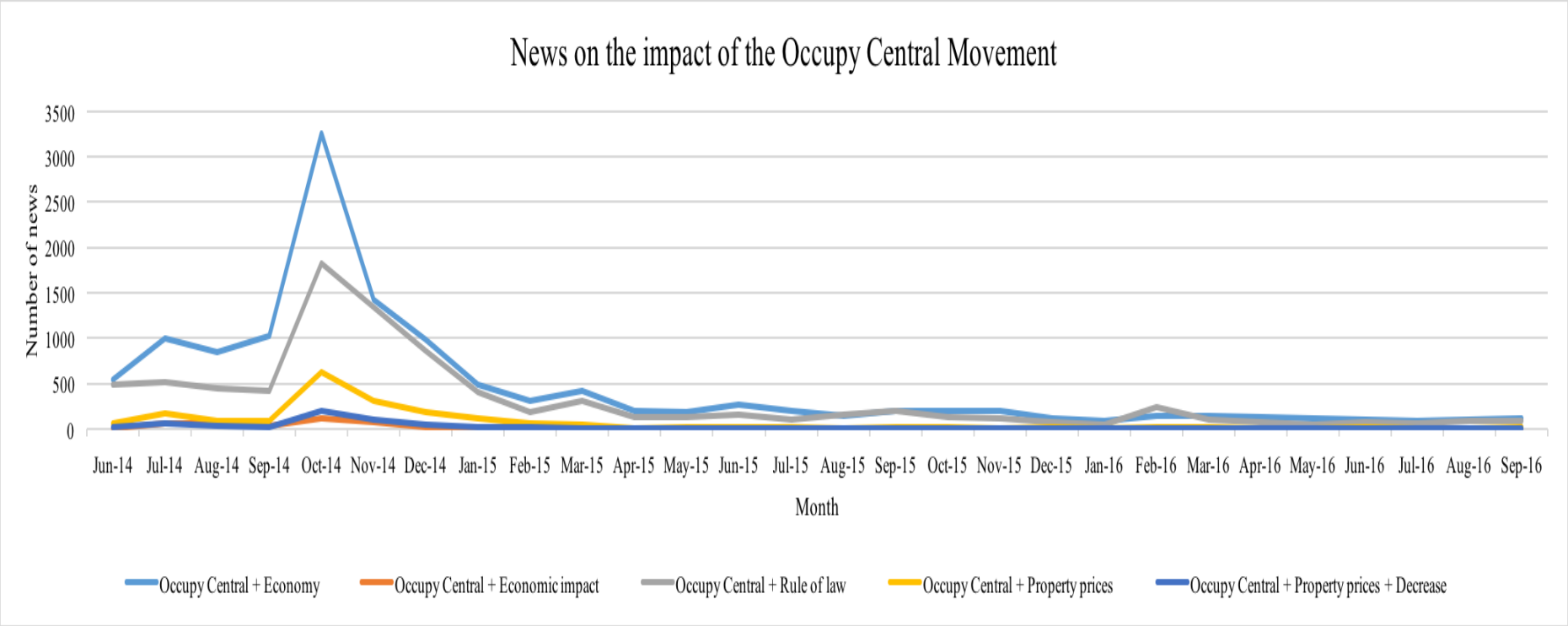
Notes: The figures show how the change in average PCA scores varies with the distance to the occupy sites. Each black spot in the graphs represents the district-level average change in a specific PCA score. The curves are fitted by LOWESS, with the gray area indicating their 95% confidence interval. The four graphs are for perceived economic insecurity, real income level reported by respondents, approval of democracy, and perceived political efficacy, respectively.

Figure A7: Attitude of Hong Kong citizens toward politicians from different camps



Note: The figure shows how the attitude of Hong Kong citizens toward politicians from different camps changes with the distance to occupy sites. It is clear that pro-Beijing and localist politicians are less popular than Pan-democratic politicians even after the protest. In addition, people who live close to the occupy sites tend to evaluate Pan-democratic politicians more highly, suggesting that the general impression of various camps is unlikely to drive our results.

Figure A8: Number of news reports on different aspects of the Umbrella Movement (2014-2016)



Note: The figure displays the time series of news reports on the Umbrella Movement that were published in the *pro-Beijing* newspapers from 2014 to 2016. The movement is called "Occupy Central" by these media. The five curves represent the monthly frequency of occurrence for different keyword combinations, marked by different colors.

Appendix C: Tables

Table A1: Administrative units in Hong Kong

	Number	Average population (2016)
Hong Kong	1	7.072 million
Legislative Council constituency	5	1,467,400
District	18	407,611
District Council constituency	431	17,023
Polling station	539	13,612

Notes: The figures are based on statistics provided by Hong Kong's Census and Statistics Department. District Council constituency is the main unit of analysis for election results. For ABS and HKES, the location of respondents is known only at district level.

Table A2: Robustness checks of the main results

	Super-seat Election			*Regular-seat Election*			
	WLS	*No re-districted*	*Public transport*	*WLS*	*No re-districted*	*Public transport*	*Lagged outcome*
Lagged vote share							0.001
							(0.026)
Driving time * Post Umbrella	0.735***	0.720***		0.315***	0.328***		0.345***
	(0.056)	(0.039)		(0.037)	(0.027)		(0.028)
Commuting time * Post Umbrella			0.275***			0.119***	
			(0.022)			(0.012)	
Post Umbrella	−3.938***	−3.928***	−2.410**	−6.654***	−7.443***	−6.410***	−8.295***
	(1.092)	(0.731)	(0.892)	(0.812)	(0.599)	(0.579)	(0.678)
Constituency FE	Y	Y	Y	Y	Y	Y	Y
N	802	766	780	802	766	780	798

Notes: The table presents results of robustness checks for the main findings. The first three columns are for the super-seat election; the next four columns are for the regular-seat election. Columns one and four display results from Weighted Least Square, where the weight is the number of voters in each constituency. Constituencies that were redistricted between 2012 and 2016 were dropped in column two and column five. In column three and six the distance to the occupation sites is calculated using public transportation time. Lagged outcome is controlled in column seven. In all the columns the coefficient of the key predictor is significant. Standard errors are clustered at constituency level. *p<0.05; **p<0.01; ***p<0.001.

Table A3: Results with covariates

	Super-seat election			*Regular-seat election*		
	With cov.	*Dist. cov.*	*LASSO*	*With cov.*	*Dist. cov.*	*LASSO*
Driving time * Post.	0.632***	0.743***	0.640***	0.175***	0.304***	0.226***
	(0.061)	(0.041)	(0.050)	(0.030)	(0.026)	(0.025)
Post Umbrella	−7.325	−7.758***	−11.333**	−1.706	−6.556***	−5.828***
	(19.958)	(1.158)	(4.074)	(12.896)	(0.849)	(1.285)
Chinese share	−1.211			−5.367		
	(12.866)			(7.899)		
Male share	3.082			−23.475		
	(33.945)			(25.006)		
Young share	33.111			35.356***		26.205**
	(16.916)			(9.893)		(8.411)
Old share	−7.315		−8.232	−10.138		
	(15.144)		(12.923)	(7.723)		
Married share	−6.346			10.142		
	(17.548)			(10.514)		
College share	22.001*			15.062*		
	(9.715)			(7.317)		
Trade share	35.212			11.199		
	(23.868)			(14.504)		
Finance share	−47.973			−68.563*		−50.124*
	(31.143)			(29.691)		(19.957)
Rich share	−14.826**			−21.542***		
	(5.277)			(4.306)		
Poor share	−16.486			1.708		
	(11.786)			(6.042)		
Mid income		0.001***			−0.0001	
		(0.0001)			(0.0001)	
Rent/Income		0.143			−0.510***	
		(0.113)			(0.098)	
With-child share			7.756			
			(5.721)			
Inc. 20k-30k share			27.288***			
			(8.119)			
Inc. 30k-40k share			18.642			
			(11.096)			
Non-Chinese share						−21.859***
						(5.462)
Public apt. share						−0.181
						(0.306)
Subsidized apt. share						0.498
						(0.340)
Constituency FE	Y	Y	Y	Y	Y	Y
N	802	802	802	802	802	802

Notes: The table presents full results of our estimation with covariates. The first three columns are for the super-seat election; the next three columns are for the regular-seat election. Columns one and three display results using covariates selected by subjective judgment (column one and three in Table 2, main text). Covariates in column three and six are chosen by LASSO (columns two and four in Table 2, main text). In columns two and four, two time-varying district characteristics are controlled. Among the covariates, "Mid income" is median income in each district, and "Rent/Income" the median value of rent over income in each district. "With-child share" is the share of families with children. "Inc. 20k-30k share" is the share of households with monthly income between 20k and 30k, and "Inc. 30k-40k share" is the share of households with monthly income between 30k and 40k. "Non Chinese share" is the share of people who are not Chinese ethnically. "Public apt. share" is the share of households living in public apartments. "Subsidized apt. share" is the share of households living in apartments subsidized by the government. Standard errors are clustered at constituency level. *p<0.05; **p<0.01; ***p<0.001.

Table A4: Results on turnout

	Vote share for the oppo.		*Turnout*	*Turnout*
	Super-seat	*Regular-seat*	*rate*	*rate*
Turnout Rate	0.991***	0.212		
	(0.226)	(0.120)		
Driving time * Post Umbrella			0.039**	
			(0.015)	
Straight-line * Post Umbrella				0.035†
				(0.018)
Post Umbrella	4.487***	−2.306**	4.327***	4.736***
	(1.190)	(0.705)	(0.321)	(0.230)

Notes: Dependent variables in the first two columns are the vote share for the opposition in the super- and regular-seat elections; the independent variable is the turnout rate in the corresponding year and constituency. In the third column turnout rate is treated as the dependent variable and the independent variable is the interaction term as in Tables 2 and 3. Standard errors are clustered at constituency level. †p<0.1; *p<0.05; **p<0.01; ***p<0.001.

Table A5: Turnout rate as a mediator

	Vote share for the oppo.	
	Super-seat	*Regular-seat*
ACME	0.026***	0.003
	[0.001, 0.067]	[−0.008, 0.014]
ADE	0.706***	0.331***
	[0.589, 0.809]	[0.266, 0.396]
Total Effect	0.731***	0.334***
	[0.619, 0.844]	[0.268, 0.400]
Prop. Mediated	0.028	0.008
	[0.002, 0.086]	[−0.026, 0.042]
N	802	802

Notes: The mediation analysis is conducted using the R package "mediation" (Imai et al., 2011). The left column is for the super-seat election, and the right column is for the regular-seat election. "ACME" is the average causal mediation effect, i.e. the effect of the interaction term on vote share through turnout rate. "ADE" is the average direct effect, the effect of the interaction term on vote share when turnout rate is hold constant. "Prop. Mediated" is the percentage of the mediation effect in the total effect. The results show that the effect of turnout rate as a mediator is very small. *p<0.05; **p<0.01; ***p<0.001.

Table A6: Mechanisms (perceived economic insecurity)

	Q9		*Q10*		*Q11*		*PCA*		*Real Income*	
With interaction?	*Yes*	*No*	*Yes*	*No*	*Yes*	*No*	*Yes*	*No*	*Yes*	*No*
Driving time * Post.	-0.022^{***}		-0.020^{**}		-0.008		-0.025^{***}		-0.009	
	(0.003)		(0.007)		(0.005)		(0.004)		(0.009)	
Post Umbrella	0.326^{***}	-0.089	0.448^{***}	0.069	0.073	-0.081	0.445^{***}	-0.030	0.587^{**}	0.428^{***}
	(0.079)	(0.057)	(0.117)	(0.072)	(0.134)	(0.061)	(0.101)	(0.071)	(0.220)	(0.117)
Gender	0.034	0.040	0.036	0.041	-0.002	0.0004	0.038	0.043	0.020	0.022
	(0.043)	(0.043)	(0.033)	(0.032)	(0.043)	(0.045)	(0.042)	(0.041)	(0.030)	(0.030)
Age	0.026^{***}	0.026^{***}	0.003	0.003	-0.015^{**}	-0.015^{**}	0.016^{**}	0.017^{**}	0.039^{***}	0.039^{***}
	(0.004)	(0.004)	(0.006)	(0.006)	(0.005)	(0.005)	(0.006)	(0.006)	(0.006)	(0.006)
Age Squared	-0.0002^{***}	-0.0002^{***}	-0.0001	-0.0001	0.0001^{**}	0.0001^{**}	-0.0002^{***}	-0.0002^{**}	-0.0003^{***}	-0.0003^{***}
	(0.00004)	(0.00004)	(0.0001)	(0.0001)	(0.00005)	(0.00005)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Married	-0.124^{**}	-0.111^{*}	-0.012	0.001	-0.052	-0.047	-0.082	-0.066	0.141	0.147
	(0.042)	(0.045)	(0.075)	(0.074)	(0.072)	(0.072)	(0.050)	(0.053)	(0.099)	(0.095)
Collage	-0.048	-0.047	-0.053	-0.052	-0.526^{***}	-0.525^{***}	-0.099^{*}	-0.094^{*}	0.889^{***}	0.890^{***}
	(0.058)	(0.056)	(0.041)	(0.045)	(0.056)	(0.056)	(0.039)	(0.039)	(0.069)	(0.068)
District FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
N	2,194	2,194	2,151	2,151	2,162	2,162	2,270	2,270	1,857	1,857

Notes: The table presents regression results on each question related to perceived economic insecurity by ABS respondents. The contents of the corresponding questions can be found in Table A10. In odd columns, both the interaction of the period dummy and the distance to the occupy sites, as well as the period dummy are added into the regression. In even columns only the period dummy is included. The last two columns show results for real income level reported by respondents (columns five and six in Table 3A, main text). Standard errors are clustered at district level. *p<0.05; **p<0.01; ***p<0.001.

Table A7: Mechanisms (perceived political efficacy)

	Q6		*Q7*		*Q8*		*PCA*	
With interaction?	*Yes*	*No*	*Yes*	*No*	*Yes*	*No*	*Yes*	*No*
Driving time * Post Umbrella	0.017		0.008		−0.003		0.010	
	(0.010)		(0.011)		(0.012)		(0.013)	
Post Umbrella	0.047	0.358***	−0.003	0.145	0.206	0.143	0.097	0.279**
	(0.218)	(0.106)	(0.211)	(0.085)	(0.234)	(0.096)	(0.250)	(0.100)
Gender	−0.252***	−0.258***	−0.114**	−0.117**	−0.230***	−0.229***	−0.255***	−0.259***
	(0.040)	(0.042)	(0.041)	(0.041)	(0.048)	(0.049)	(0.045)	(0.047)
Age	−0.012	−0.012	−0.004	−0.004	−0.007	−0.007	−0.010	−0.010
	(0.009)	(0.009)	(0.007)	(0.007)	(0.006)	(0.006)	(0.008)	(0.008)
Age Squared	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.00005)	(0.00005)	(0.0001)	(0.0001)
Married	−0.052	−0.062	−0.064	−0.069	0.089	0.090	−0.007	−0.012
	(0.062)	(0.064)	(0.050)	(0.050)	(0.054)	(0.054)	(0.061)	(0.062)
Collage	0.475***	0.476***	0.361***	0.361***	0.347***	0.347***	0.517***	0.517***
	(0.069)	(0.069)	(0.051)	(0.052)	(0.068)	(0.068)	(0.070)	(0.070)
District FE	Y	Y	Y	Y	Y	Y	Y	Y
N	2,128	2,128	2,121	2,121	2,091	2,091	2,035	2,035

Notes: The table presents regression results on each question related to the level of perceived political efficacy felt by ABS respondents. The contents of the corresponding questions can be found in Table A10. In odd columns, both the interaction of the period dummy and the distance to the protest sites, as well as the period dummy are added into the regression. In even columns only the period dummy is included. Standard errors are clustered at district level. *p<0.05; **p<0.01; ***p<0.001.

Table A8: Mechanisms (approval of democracy)

	Q1		*Q2*		*Q3*	
With interaction?	*Yes*	*No*	*Yes*	*No*	*Yes*	*No*
Driving time * Post Umbrella	−0.003		0.001		−0.009	
	(0.012)		(0.012)		(0.007)	
Post Umbrella	0.027	−0.029	−0.669*	−0.653***	0.158	−0.001
	(0.225)	(0.120)	(0.272)	(0.146)	(0.243)	(0.132)
Gender	−0.084*	−0.082*	0.136***	0.135***	0.120**	0.124***
	(0.040)	(0.040)	(0.033)	(0.034)	(0.037)	(0.036)
Age	−0.012	−0.012	−0.0004	−0.0004	−0.010	−0.010
	(0.008)	(0.008)	(0.009)	(0.009)	(0.008)	(0.008)
Age Squared	0.0001	0.0001	0.00000	0.00000	0.0001	0.0001
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Married	−0.051	−0.049	0.069	0.068	−0.019	−0.013
	(0.070)	(0.072)	(0.059)	(0.060)	(0.063)	(0.065)
College	−0.085	−0.085	−0.019	−0.019	0.002	0.002
	(0.063)	(0.064)	(0.068)	(0.067)	(0.075)	(0.076)
District FE	Y	Y	Y	Y	Y	Y
N	1,865	1,865	1,911	1,911	2,042	2,042
	Q4		*Q5*		*PCA*	
With Interaction?	*Yes*	*No*	*Yes*	*No*	*Yes*	*No*
Driving time * Post Umbrella	−0.005		−0.002		−0.003	
	(0.011)		(0.009)		(0.013)	
Post Umbrella	−0.011	−0.100	0.158	0.117	−0.201	−0.248
	(0.257)	(0.114)	(0.201)	(0.097)	(0.342)	(0.189)
Gender	0.049	0.050	−0.062*	−0.061	0.092	0.093
	(0.057)	(0.058)	(0.032)	(0.032)	(0.051)	(0.054)
Age	−0.011*	−0.011*	−0.030***	−0.030***	−0.027**	−0.027**
	(0.005)	(0.005)	(0.008)	(0.008)	(0.009)	(0.009)
Age Squared	0.0001	0.0001	0.0003***	0.0003***	0.0002*	0.0002*
	(0.00004)	(0.00004)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Married	0.122*	0.124*	−0.174**	−0.173**	−0.018	−0.017
	(0.049)	(0.053)	(0.058)	(0.058)	(0.094)	(0.098)
College	0.029	0.029	0.124*	0.124*	−0.044	−0.045
	(0.052)	(0.052)	(0.058)	(0.058)	(0.092)	(0.091)
District FE	Y	Y	Y	Y	Y	Y
N	1,849	1,849	2,167	2,167	1,476	1,476

Notes: The table presents regression results on each question related to ABS respondents' approval for democracy. The contents of the corresponding questions can be found in Table A10. In odd columns, both the interaction of the period dummy and the distance to the occupation sites, as well as the period dummy are added into the regression. In even columns only the period dummy is included. Standard errors are clustered at district level. *p<0.05; **p<0.01; ***p<0.001.

Table A10: ABS questions related to mechanisms

Q1	Democracy may have its problems, but it is still the best form of government	1 Strongly agree
		2 Agree
		3 Disagree
		4 Strongly disagree
		8 Can't choose
		9 Decline to answer
Q2	Where would you want our country to be democratic in the future?	1 Complete Undemocratic
		10 Complete Democratic
		77 Do not understand the question
		78 Can't choose
		99 Decline to answer
Q3	Which would you think democracy is suitable for our country?	1 Democracy is completely unsuitable
		10 Democracy is perfectly suitable
		77 Do not understand the question
		78 Can't choose
		99 Decline to answer
Q4	Which of the following statements comes closest to your own opinion?	1 Democracy is always preferable to any other kind of government
		2 Under some circumstances, an authoritarian government can be preferable to a democratic one
		3 For people like me, it does not matter whether we have a democratic or a nondemocratic regime
		8 Can't choose
		9 Decline to answer
Q5	If you had to choose between democracy and economic development, which would you say is more important?	1 Economic development is definitely more important
		2 Economic development is somewhat more important
		3 Democracy is somewhat more important
		4 Democracy is definitely more important
		5 They are both equally important
		8 Can't choose
		9 Decline to answer
Q6	I think I have the ability to participate in politics.	1 Strongly agree
		2 Somewhat agree
		3 Somewhat disagree
		4 Strongly disagree
		8 Can't choose
		9 Decline to answer
Q7	Sometimes politics and government seems so complicated that a person like me can't really understand what is going on.	1 Strongly agree
		2 Somewhat agree
		3 Somewhat disagree
		4 Strongly disagree
		8 Can't choose
		9 Decline to answer
Q8	People like me don't have any influence over what the government does.	1 Strongly agree
		2 Somewhat agree
		3 Somewhat disagree
		4 Strongly disagree
		8 Can't choose
		9 Decline to answer
Q9	How would you rate the overall economic condition of our country today?	1 Very good
		2 Good
		3 So so (not good nor bad)
		4 Bad
		5 Very bad
		8 Can't choose
		9 Decline to answer
Q10	How would you describe the change in the economic condition of our country over the last years?	1 Much better
		2 A little worse
		3 About the same
		4 A little worse
		5 Much worse
		8 Can't choose
		9 Decline to answer
Q11	Does the total income of your household allow you to satisfactorily cover tour needs?	1 Our income covers the needs well, we can save.
		2 Our income covers the needs all right, without much difficulties
		3 Our income does not cover the needs, there are difficulties
		4 Our income does not cover the needs, there are great difficulties
		7 Do not understand the question
		9 Decline to answer

Notes: The table displays all the questions in the ABS dataset we use to construct the PCA scores. The number of the questions corresponds to those appearing in regression tables above. Answers such as "Can't choose" and "Decline to answer" are dropped in the analysis. And all the answers are recoded so that larger values imply that one is more positive along the particular dimension.

Table A9: Conversion matrix

	Pro-establishment 2016	Pan-democratic 2016	Localist 2016	Total 2016
Pro-establishment 2012	260	38	19	317
Pan-democratic 2012	79	383	166	628
Total 2012	339	421	185	945

Notes: The figures are based on HKES data. Each row shows how many supporters of a camp in 2012 converted to other camps in 2016. There are only two rows as the *localist camp* didn't exist in 2012.