Rules for Recovery: Impact of Indexed Disaster Funds on Shock Coping in Mexico - Online Appendix

By: Alejandro del Valle, Alain de Janvry and Elisabeth Sadoulet

A Predicting State level GDP growth with night lights

While our primary interest lies in determining whether night lights growth can predict municipal GDP growth, in the absence of this data, we instead investigate the relationship between the growth of night lights and state GDP growth. Given that our primary purpose is to estimate the inverse elasticity of light to GDP that would allow us to back out the impact of Fonden on local economic activity, we restrict the sample to the 28 states that have requested Fonden funding in the 2004-2012 period. We follow the approach taken by Henderson, Storeygard and Weil (2011) and focus primarily on determining whether night lights can predict year-to-year growth, annual fluctuations, recession and expansions, and long-term growth.

Table A1 presents the main results. The baseline specification regresses log GDP on log night lights, state fixed effects, and year fixed effects. Standard errors are clustered at the state level in all cases. Column 1 reports the results of this exercise, where we sharply estimate an elasticity of roughly 0.21.

Next, in column 2, we test whether night lights are capable of predicting annual fluctuations by extending the previous specification to include state trends. Since we are primarily interested in short term deviations from the state growth path, this is the key specification for our analysis. As in the previous case, we sharply estimate an elasticity in the order of 0.1 (p-val=0.02).[1] This result is important because it suggests that night lights do a reasonably good job of predicting annual fluctuations in GDP.

In column 3, we test for ratchet effects, that is, whether relative to the state mean over time, increases and decreases in night lights are symmetrically related to increases and decreases in GDP. This calculation is performed in two steps: (i) We demean the data by regressing GDP and lights on year and state fixed effects. (ii) We regress the GDP residuals on absolute value positive lights residuals and absolute value negative lights residuals. We find that the coefficients are very similar in magnitude and that they have the opposite signs. We thus conclude that night lights are capable of picking up both economic expansions and

[1]The wild cluster bootstrap (5000 iterations) produces a p-value of 0.027

economic downturns.

In terms of R^2, note that the R^2 reported in columns 1 and 2 is a within state R^2, and thus it still accounts for the role of year dummies. The R^2 reported in column 3, 8%, reflects the contribution of night lights to explaining within-state and within-year variation in GDP solely.

Last, in column 4, we test whether night lights can be used to predict long-term growth. This test is performed using the long difference specification where we regress the change in log GDP between 2004 and 2013 on the change in log night lights between 2004 and 2013. We find a positive and sharply estimated elasticity, and an R^2 of 0.38. All in all, while our sample size is small compared to those of other papers in the literature, our results validate the idea of using night lights as a proxy for economic activity at the subnational level in Mexico.

Table A1: Elasticity of night lights to GDP

	Base specification	Annual fluctuations	Asymmetric fluctuations	Long difference
Dependent Variable:	ln(GDP)	ln(GDP)	ln(GDP) residuals	ln(GDP)
	(1)	(2)	(3)	(4)
ln(lights/area)	0.215	0.095		0.618
	(0.065)	(0.038)		(0.163)
	[0.003]	[0.020]		[0.001]
|*plus ln(lights/area) residuals* |			0.169	
			(0.118)	
			[0.162]	
|*minus ln(lights/area) residuals* |			-0.256	
			(0.111)	
			[0.029]	
R-squared	0.89	0.96	0.08	0.38
Observations	280	280	280	28
State FE	yes	yes	In demean	-
Year FE	yes	yes	In demean	-
State Trend	-	yes	-	-

Note: Standard errors clustered at the state level are in parentheses, p-values are in squared brackets. Mexico has 32 states, the sample has been restricted to the 28 states that received Fonden between 2004 and 2012.

B Additional Tables and Figures

Table A2: Complier probability derivative and treatment effect derivative

	(1)	(2)
Complier Probability Derivative	0.0020	0.0002
p-value	0.640	0.983
Treatment Effect Derivative	0.0028	0.0028
p-value	0.802	0.882
Bandwidth (mm)	74.3	48.7
Obs (left|right)	1287|605	894|460

Note: This table provides estimates of the complier probability derivate as described in Cerulli et al. (2017), and of the treatment effect derivative as described in Dong and Lewbel (2015). All specifications use a triangular kernel and a local quadratic polynomial. The bandwidth selection algorithm used in column 1 is optimal for point estimation; the selection algorithm in column 2 is optimal for inference of confidence intervals.

Table A3: Robustness Donut-hole analysis

Donut-hole Radius	RD Estimate	p-value	CI 95%	Bandwidth	Obs	Excluded obs Left	Excluded obs Right
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
0.0	0.260	0.009	[0.08,0.56]	57.9	1563	0	0
0.5	0.283	0.008	[0.09,0.60]	58.3	1569	2	7
1.0	0.311	0.006	[0.12,0.73]	66.8	1725	27	10
1.5	0.412	0.011	[0.13,0.99]	51.3	1376	31	13
2.0	0.470	0.007	[0.18,1.17]	56.3	1467	37	21
2.5	0.312	0.037	[0.03,0.83]	53.1	1358	59	26

Note: Robustness of the LATE estimate to the exclusion of observations that are within 2.5 mm on each side of the threshold. The dependent variable is the log difference night lights between the 12 months before and after a disaster. Column 1 reports the excluded radius in mm, columns 7 and 8 report the number of observations excluded at each side of the threshold. Estimates are derived using a triangular kernel, a local linear polynomial, and a h_{MSE} optimal bandwidth. The p-values and 95% confidence intervals reported are constructed using robust bias correction and clustering at the municipal level.

Table A4: Placebo cutoffs

Alternative cutoff (1)	RD Estimate (2)	p-value (3)	CI 95% (4)	Bandwidth (5)	Obs Left (6)	Obs Right (7)
-46.2	0.329	0.663	[-0.8,1.3]	19.1	312	348
-36.3	-1.464	0.928	[-13.2,14.5]	16.3	257	341
-27.1	0.583	0.748	[-4.5,6.3]	10.3	193	216
-19.7	-0.328	0.388	[-1.1,0.4]	8.0	188	120
-9.2	-0.031	0.797	[-0.3,0.3]	6.5	113	111
0.0	0.260	0.015	[0.1,0.6]	58.7	1052	529
8.0	-1.144	0.427	[-3.6,1.5]	21.6	98	243
15.3	-0.144	0.393	[-1.3,0.5]	15.8	196	153
26.2	6.195	0.676	[-133.1,205.2]	16.6	162	118
37.6	0.784	0.628	[-3.4,5.7]	13.2	109	86
52.2	4.517	0.188	[-9.5,48.4]	16.8	111	86

Note: The dependent variable is the log difference night lights between the 12 months before and after a disaster. Estimates of the LATE at the true zero cut-off and at various placebo cut-offs. The sample in the last five rows is restricted to nonnegative values of the running variable, the placebo cut-off are given by the first five deciles. The sample in the first 5 rows is restricted to negative values of the running variable, the placebo cut-offs are determined in an analogous manner. Estimates are derived using a triangular kernel, a local linear polynomial, and a h_{MSE} optimal bandwidth. The p-values and 95% confidence intervals reported are constructed using robust bias correction and clustering at the municipal level.

Table A5: Impact of Fonden on night lights, robustness (tuning parameters)

	Local Polynomial Degree		Kernel		Alternative bandwidths		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
LATE (τ_{FRD})	0.312	0.337	0.218	0.252	0.274	0.258	0.292
p-value	0.051	0.024	0.002	0.011	0.020	0.014	0.024
CI 95%	[-0.00,0.64]	[0.04,0.63]	[0.10,0.41]	[0.07,0.55]	[0.05,0.57]	[0.06,0.49]	[0.04,0.56]
Bandwidth (left|right) (mm)	74.3|74.3	48.7|48.7	62.1|62.1	54.3|54.3	50.1|50.1	42.0|71.3	29.0|49.2
Obs (left|right)	1287|605	894|460	1113|549	979|509	926|478	780|599	547|466
Bandwidth selection	$\hat{h}_{MSE}$	$\hat{h}_{CER}$	$\hat{h}_{MSE}$	$\hat{h}_{MSE}$	ITT $\hat{h}_{MSE}$	$\hat{h}_{MSE2}$	$\hat{h}_{CER2}$
Local Polynomial Degree	2	2	1	1	1	1	1

Note: The dependent variable is the log difference night lights between the 12 months before and after a disaster. Estimates are derived using a triangular kernel with the exception of columns 3 (epanechnikov) and 4 (uniform). The local polynomial degree, and optimal bandwidth selection algorithm is indicated in each column. The $\hat{h}_{MSE}$ bandwidth selection algorithm is optimal for point estimation; the $\hat{h}_{CER}$ selection algorithm is optimal for inference of confidence intervals. The subscript 2 in the description of the bandwidth selection algorithm denotes that different bandwidth lengths have been selected in each side of the threshold. The p-values and 95% confidence intervals reported are constructed using robust bias correction and clustering at the municipal level.

Table A6: Impact of Fonden on night lights, robustness checks

	Night lights alternative definitions			Multiple exposure to Fonden		Extreme thresholds	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
LATE (τ_{FRD})	0.273	0.262	0.156	0.257	0.238	0.253	0.229
p-value	0.011	0.009	0.068	0.015	0.026	0.020	0.018
CI 95%	[0.07,0.52]	[0.08,0.56]	[-0.01,0.39]	[0.06,0.54]	[0.03,0.52]	[0.05,0.56]	[0.05,0.54]
Bandwidth (mm)	49.9	61.2	63.1	49.7	56.1	55.8	52.7
Obs (left|right)	918|467	1094|533	1132|552	695|338	529|263	924|440	841|450

Note: The dependent variable is the log difference night lights between the 12 months before and after a disaster. In column 1, the log difference night lights is constructed using only information form the newest satellite available. In column 2, the log difference night lights excludes top coded pixels ($DN \geq 55$). The specification in column 3 includes year fixed effects. Columns 4, excludes observations with Fonden in the year before or after. Column 5, excludes observations with Fonden two years before or after. Column 6, excludes observations in the bottom decile of Fonden thresholds. Column 7, excludes observations in the top decile of Fonden thresholds. All estimates are derived using a triangular kernel, a local linear polynomial, and a h_{MSE} optimal bandwidth. The p-values and 95% confidence intervals reported are constructed using robust bias correction and clustering at the municipal level.

Table A7: Additional datasets and Sources

Variable	**Source**
INEGI (The National Institute of Statistics and Geography)	
Population	
No. of dwellings	
Pop. with no social security	
Pop. 15 or older illiterate	Census
Pop. 15 or older with no schooling	INEGI (2000)
Dwellings with electricity	INEGI (2005)
Dwellings with piped water	INEGI (2010)
Dwellings connected to sewage	
Dwellings with a refrigerator	
Elevation	
Revenue of municipal government	Public finances of municipalities
Expenditures of municipal government	INEGI (2014)
Total transfers	
Revenue sharing transfers	
Conditional transfers	
Municipal surface area	Municipal boundaries
Centroid longitude	INEGI (2013*a*)
Centroid latitude	
State level GDP	INEGI (2013*b*)
Fund for Natural Disasters Fonden	
Fonden expenditures	Fonden administrative record
Fonden disbursement times	Fonden (2015)
Fonden planned reconstruction times	
Other Datasets	
Historic mean annual rainfall (20 years before Fonden)	Conagua (2015)
Health units per 100,000	DGIS (2001)
Infant mortality rate	Conapo (2005)
Road intersection density (per 100 sq. km)	USGS (2003)
Road density (per 100 sq. km)	
GDP deflator	World Bank (2010*a*)
PPP exchange rates	World Bank (2010*b*)

Note: The terms of use of the INEGI datasets used can be found in `http://en.www.inegi.org.mx/inegi/terminos.html`

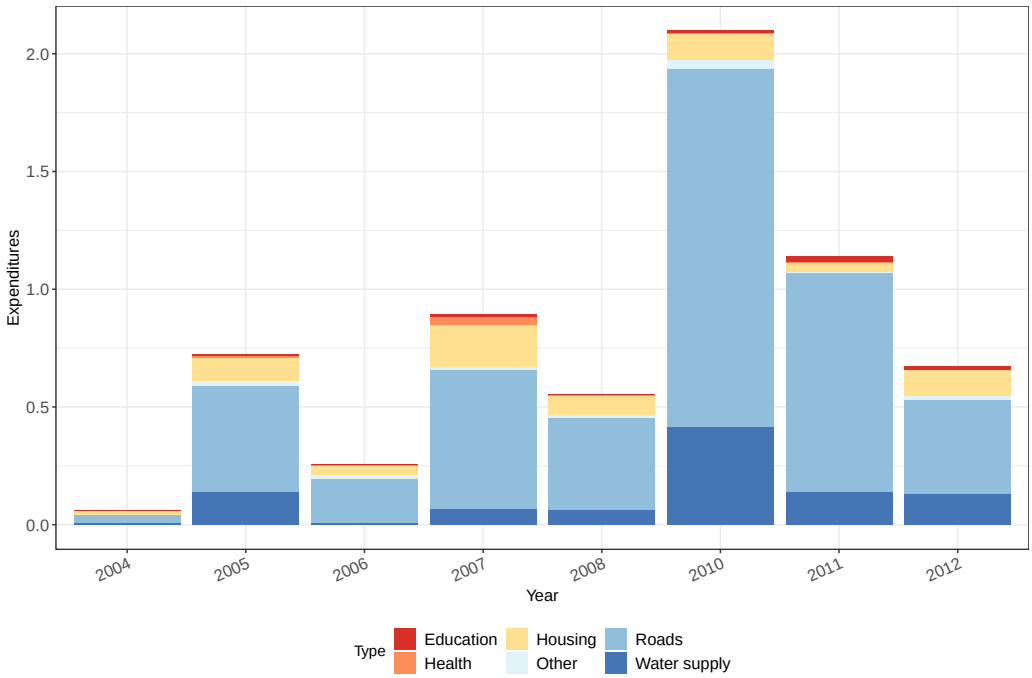


Figure A1: Fonden expenditures by year and type of reconstruction

Note: Monetary values are in billion constant 2010 international dollars. Source: Fonden (2015)

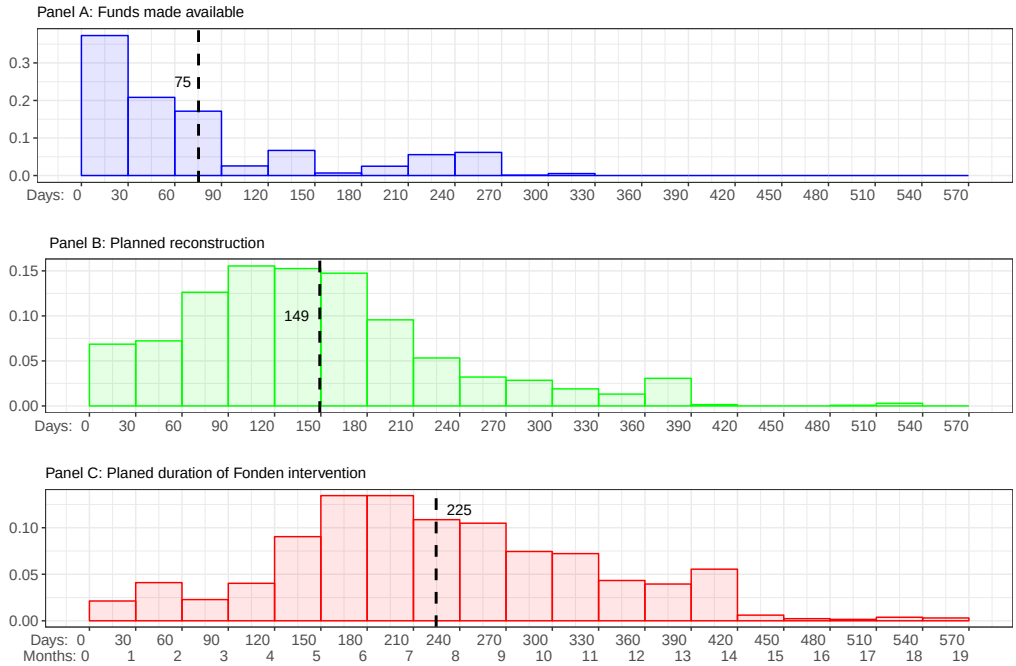


Figure A2: Fonden fund disbursement and planned reconstruction times

Note: The dashed line in each panel represents the average number of days. Panel B plots planned and not realized reconstruction days. Panel C is the sum of days reported in panel A and B. Source: Fonden (2015)

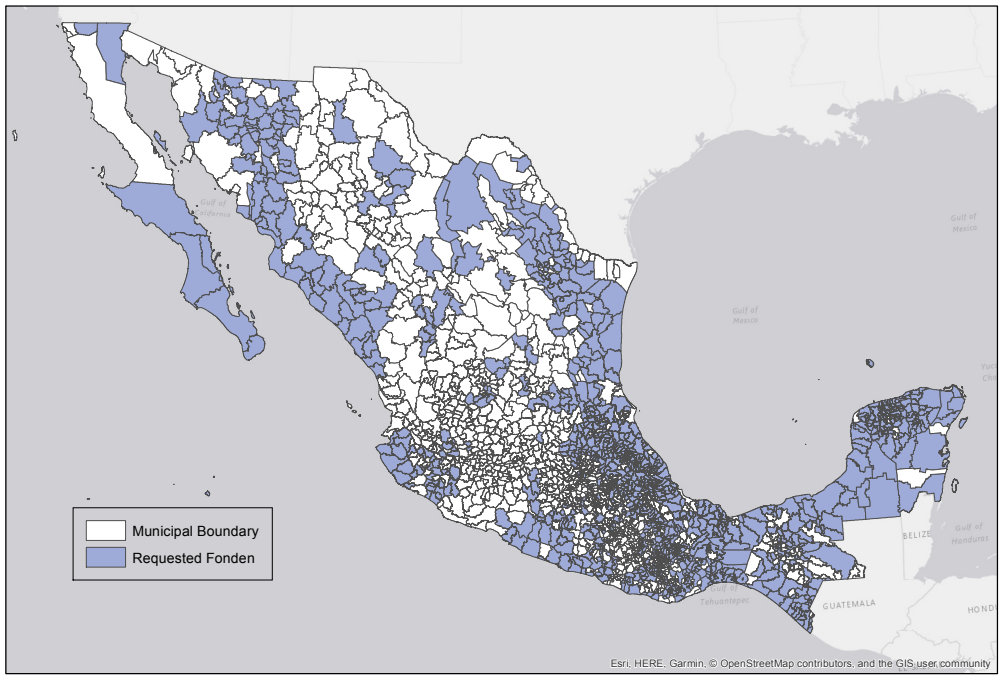


Figure A3: Municipalities that requested verification for Fonden

Note: Map of Mexico, municipal boundaries are delineated by thin grey lines. During our sample period, 1198 municipalities requested Fonden. Source: INEGI (2013*a*); Boudreau (2015)

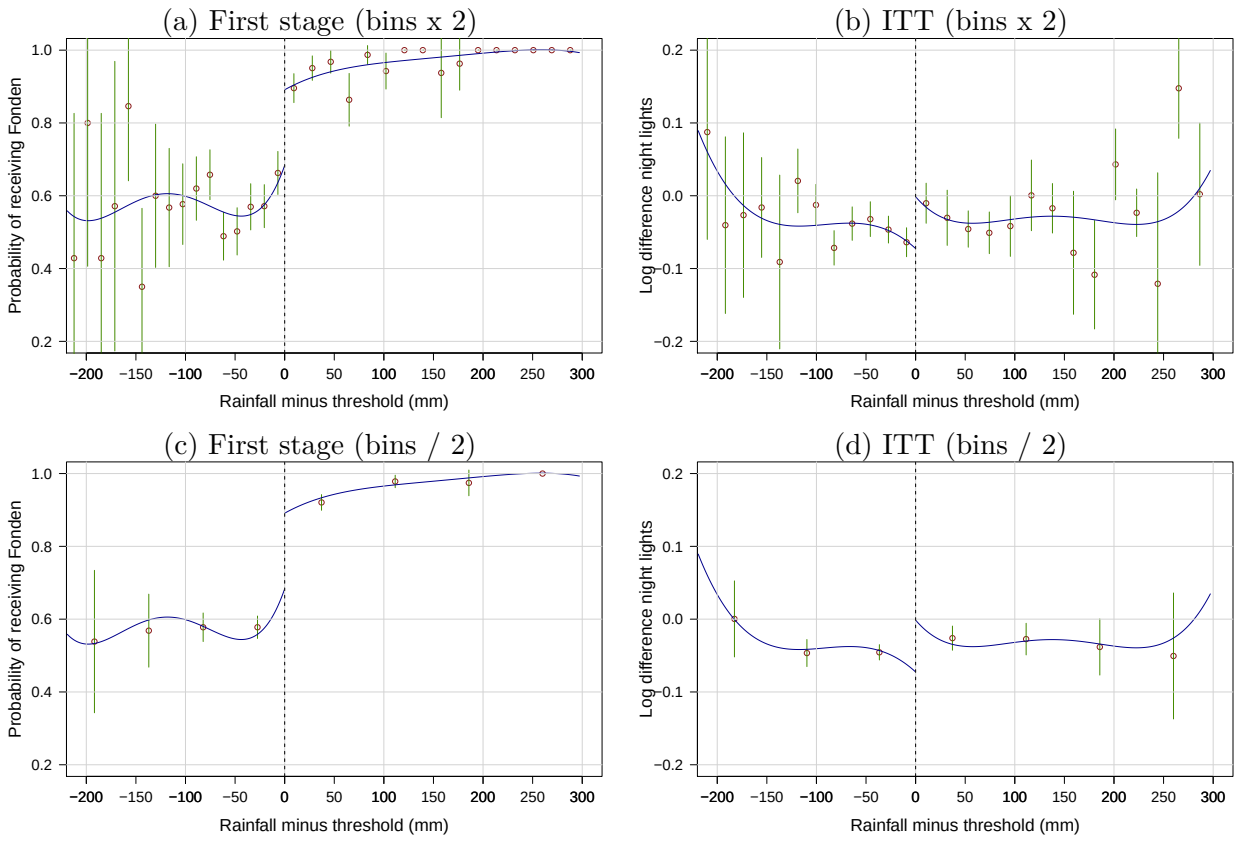


Figure A4: First stage and Intention-to-Treat, robustness

Note: Graphs A4a and A4c plots the probability of receiving Fonden as a function of the running variable (rainfall minus threshold). Graphs A4b and A4d plots the log difference night lights between 12 months before and after a disaster as a function of the running variable. In each graph, the support of the running variable has been partitioned into disjoint bins. The number of bins is selected to minimize the integrated mean square error of the underlying regression function, as described in Calonico, Cattaneo and Titiunik (2015). In graphs A4a and A4b the number of bins has been doubled. In graphs A4c and A4d the number of bins has been halved. The circles plot the local mean of the outcome at the mid-point of each bin. The error bars are the 95% confidence intervals for the local means. The solid lines are fourth-order global polynomials fits (estimated separately on each side of the threshold). Observations to the right of the vertical dashed line are eligible for Fonden under the heavy rainfall criteria.

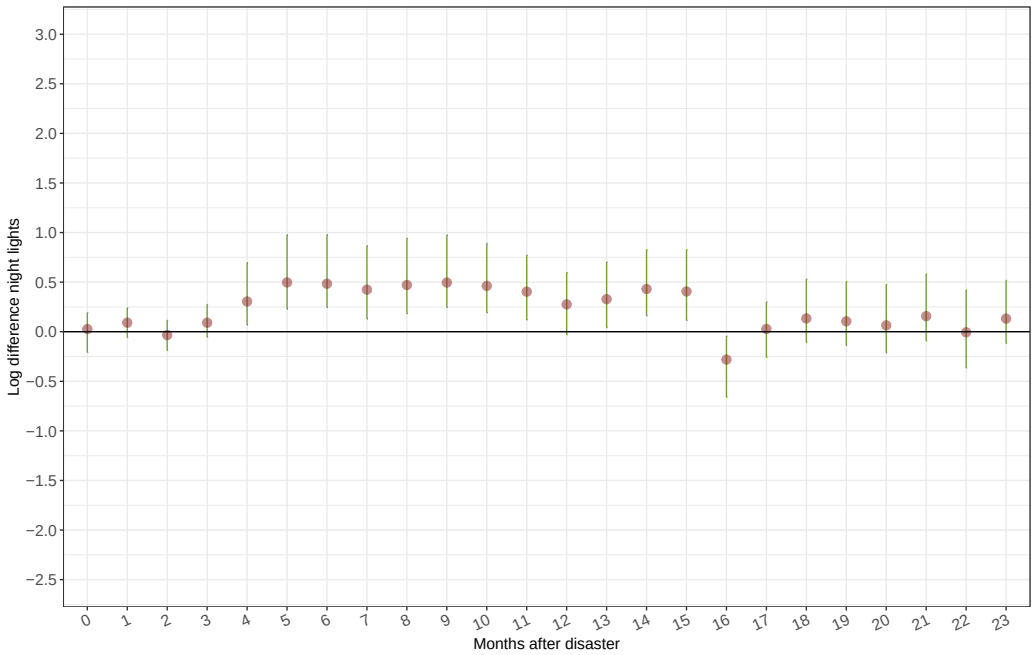


Figure A5: Month by month impact of Fonden after a disaster (interpolated dataset)

Note: This figure plots coefficients and robust 95% confidence intervals of Fonden LATE. Missing values in the night lights dataset, used to construct the 24 outcome variables, has been filled using pixel level linear interpolation. The outcome variables are the log difference in night lights between the average 12 months before a disaster (months -12 to -1) and each month in the two years after a disaster. Each plotted coefficient is estimated independently using a triangular kernel, a local linear polynomial, and the average h_{MSE} bandwidth.

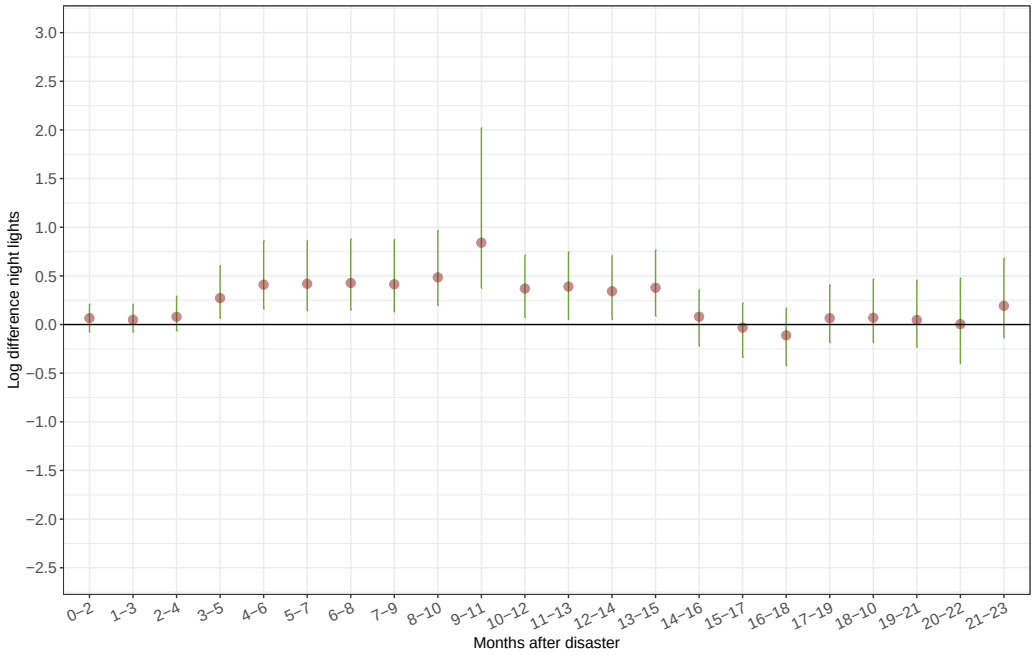
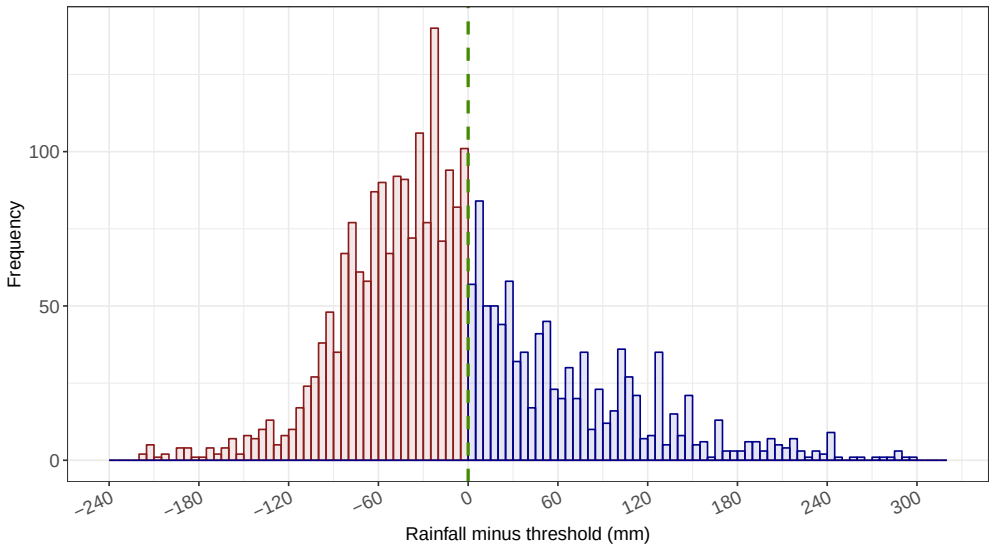


Figure A6: Month by month impact of Fonden after a disaster (3 month averages)

Note: This figure plots coefficients and robust 95% confidence intervals of Fonden LATE. The 22 outcome variables are the log difference in night lights between the 12 months before a disaster (months -12 to -1) and the three month average in the post-disaster period indicated in the axis label. Each plotted coefficient is estimated independently using a triangular kernel, a local linear polynomial, and the average h_{MSE} bandwidth.

(a) Histogram



(b) Estimated Density

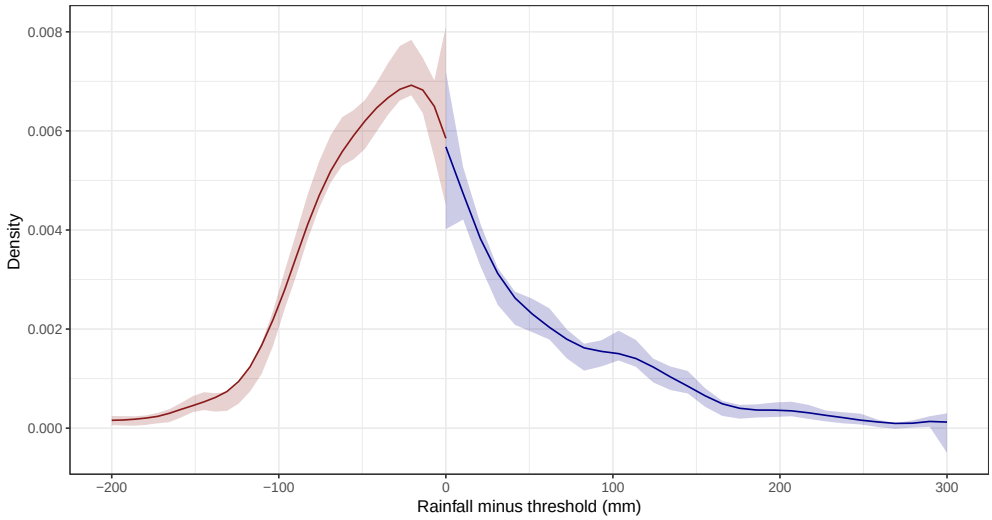


Figure A7: Histogram and estimated density of rainfall minus threshold (full support)

Note: Panel A and B plot the histogram and empirical density of the running variable (rainfall minus threshold) for the entire support of the running variable. The figure in Panel B is derived using the methods proposed by Cattaneo, Jansson and Ma (2018).

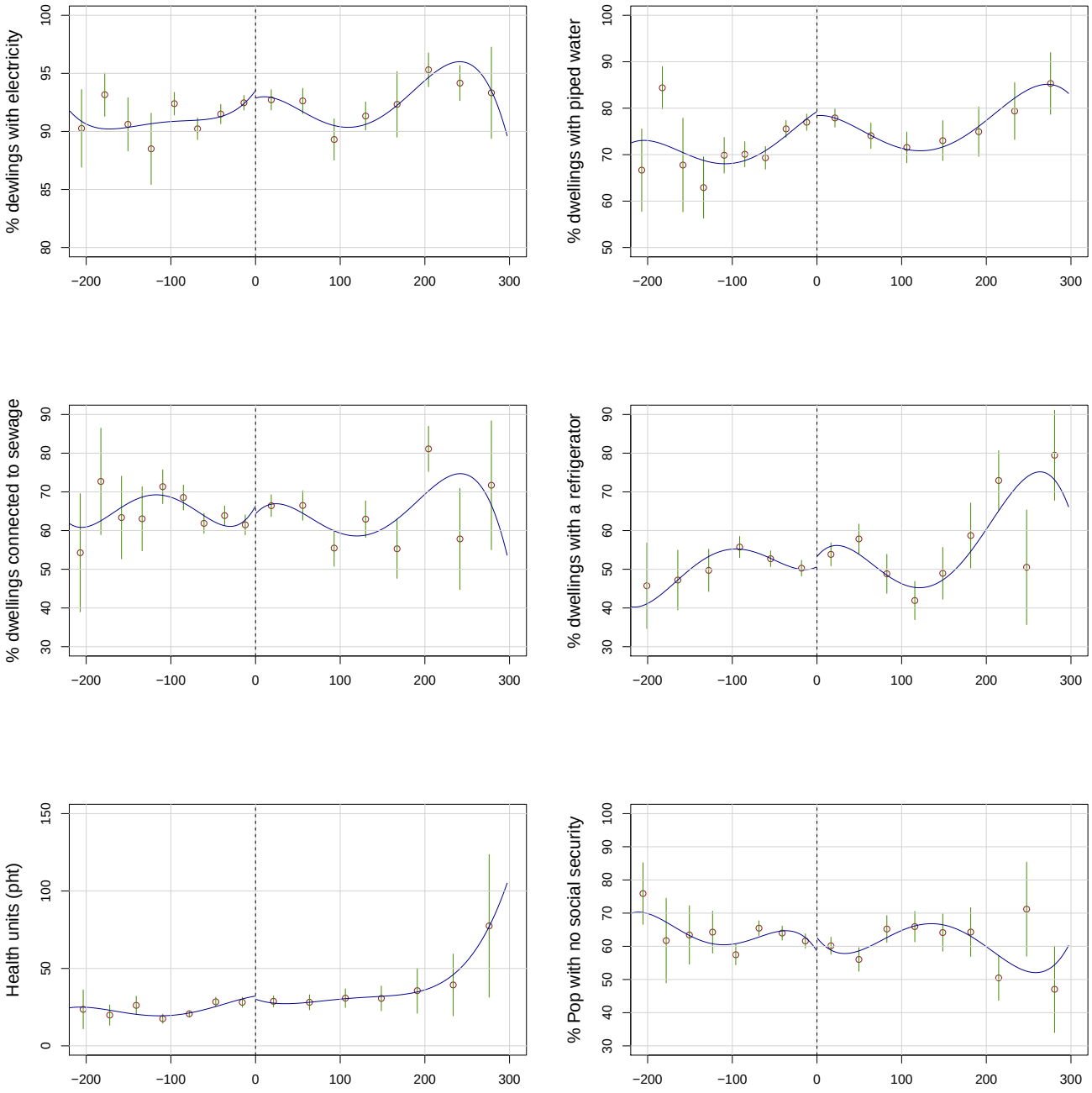


Figure A8: Predetermined Covariates I

Note: Each graph plots the outcome listed on the label as a function of the running variable (rainfall minus threshold). In each graph, the support of the running variable has been partitioned into disjoint bins. The number of bins is selected to minimize the integrated mean square error of the underlying regression function, as described in Calonico, Cattaneo and Titiunik (2015). The circles plot the local mean of the outcome at the mid-point of each bin. The error bars are the 95% confidence intervals for the local means. The solid lines are fourth-order global polynomials fits (estimated separately on each side of the threshold). Observations to the right of the vertical dashed line are eligible for Fonden under the heavy rainfall criteria.

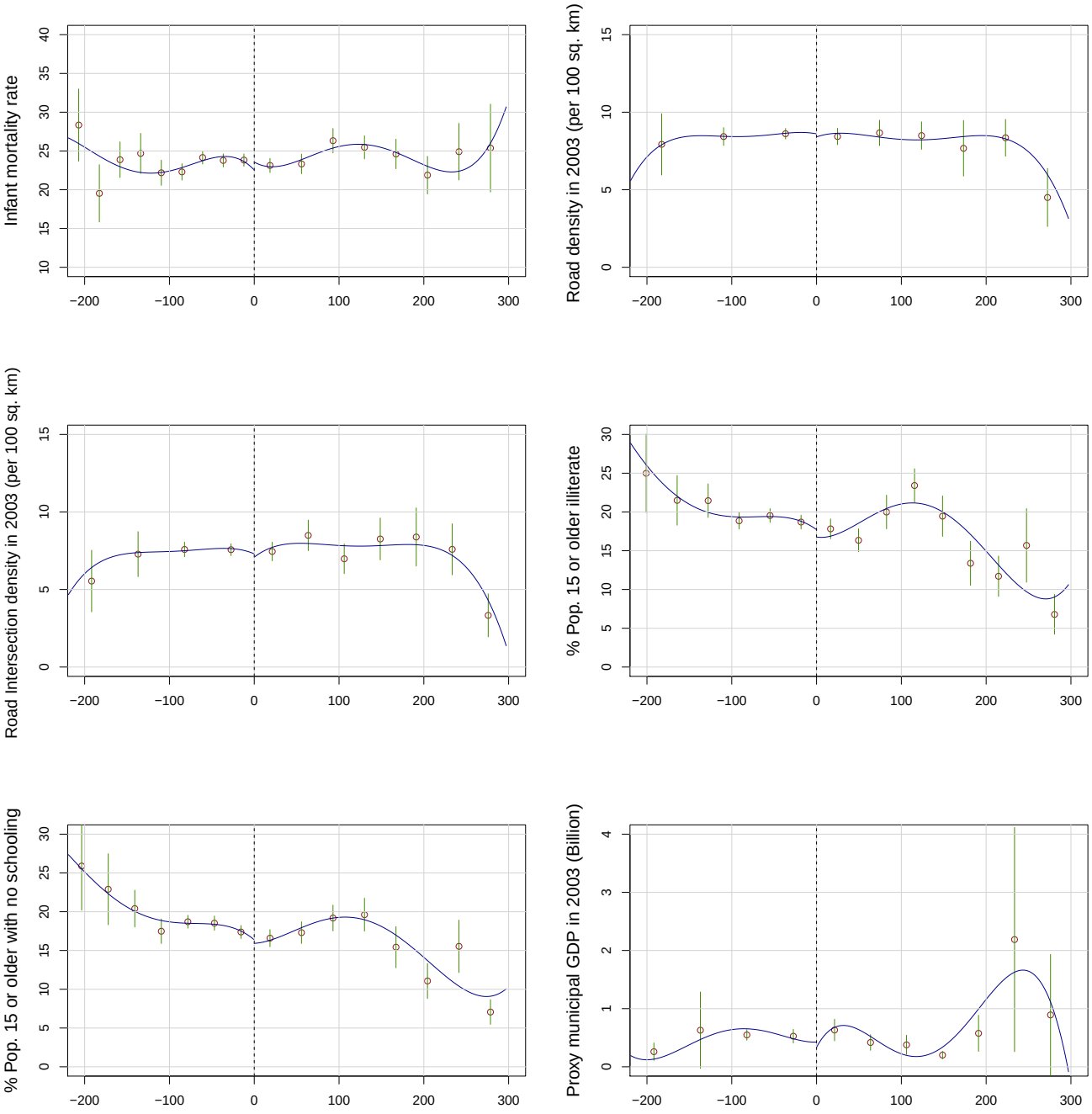


Figure A9: Predetermined Covariates II

Note: Each graph plots the outcome listed on the label as a function of the running variable (rainfall minus threshold). In each graph, the support of the running variable has been partitioned into disjoint bins. The number of bins is selected to minimize the integrated mean square error of the underlying regression function, as described in Calonico, Cattaneo and Titiunik (2015). The circles plot the local mean of the outcome at the mid-point of each bin. The error bars are the 95% confidence intervals for the local means. The solid lines are fourth-order global polynomials fits (estimated separately on each side of the threshold). Observations to the right of the vertical dashed line are eligible for Fonden under the heavy rainfall criteria.

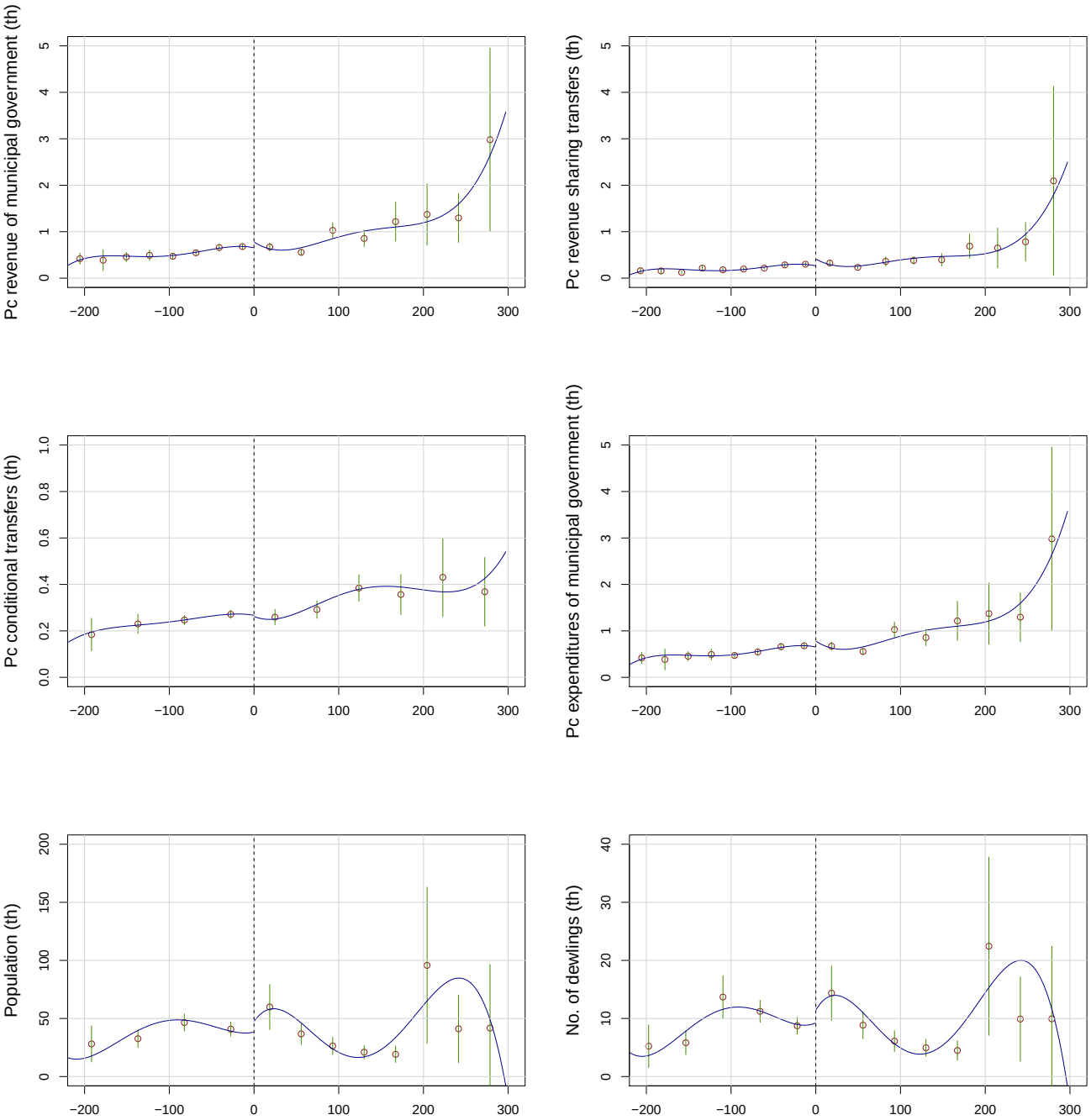


Figure A10: Predetermined Covariates III

Note: The abbreviation (pc) denotes per capita, (th) denotes thousand. Each graph plots the outcome listed on the label as a function of the running variable (rainfall minus threshold). In each graph, the support of the running variable has been partitioned into disjoint bins. The number of bins is selected to minimize the integrated mean square error of the underlying regression function, as described in Calonico, Cattaneo and Titiunik (2015). The circles plot the local mean of the outcome at the mid-point of each bin. The error bars are the 95% confidence intervals for the local means. The solid lines are fourth-order global polynomials fits (estimated separately on each side of the threshold). Observations to the right of the vertical dashed line are eligible for Fonden under the heavy rainfall criteria.

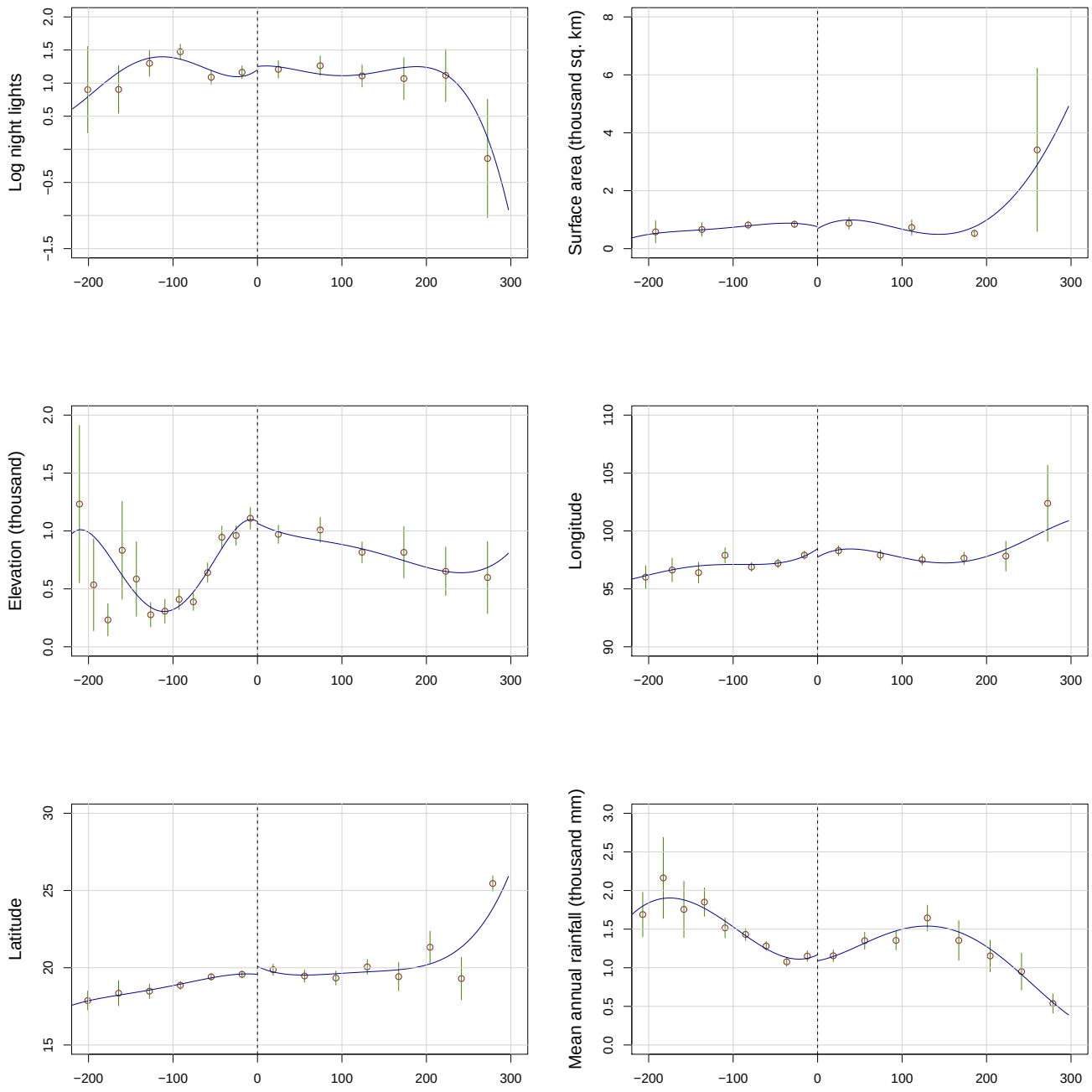


Figure A11: Predetermined Covariates IV

Note: Each graph plots the outcome listed on the label as a function of the running variable (rainfall minus threshold). In each graph, the support of the running variable has been partitioned into disjoint bins. The number of bins is selected to minimize the integrated mean square error of the underlying regression function, as described in Calonico, Cattaneo and Titiunik (2015). The circles plot the local mean of the outcome at the mid-point of each bin. The error bars are the 95% confidence intervals for the local means. The solid lines are fourth-order global polynomials fits (estimated separately on each side of the threshold). Observations to the right of the vertical dashed line are eligible for Fonden under the heavy rainfall criteria.

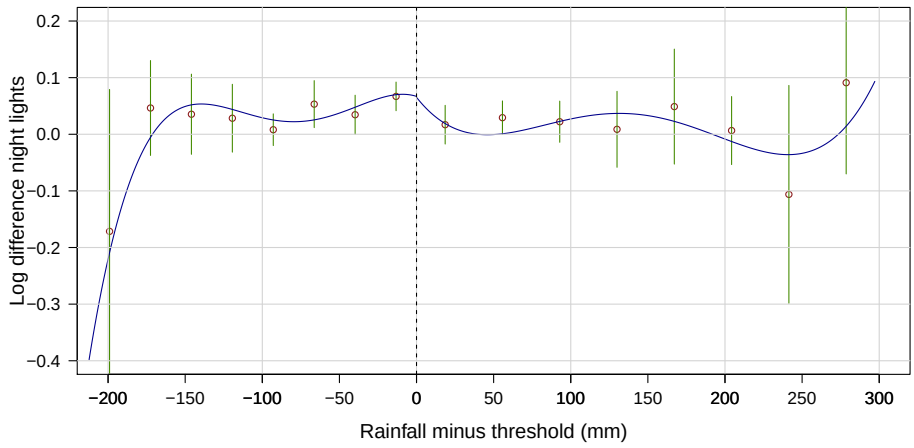


Figure A12: Intention-to-treat (placebo)

Note: The figure plots the log difference night lights, between two years before an event (months -24 to -13) and the year before (months -12 to -1), as a function of the running variable (rainfall minus threshold). The support of the running variable has been partitioned into disjoint bins. The number of bins is selected to minimize the integrated mean square error of the underlying regression function, as described in Calonico, Cattaneo and Titiunik (2015). The circles plot the local mean of the outcome at the mid-point of each bin. The error bars are the 95% confidence intervals for the local means. The solid lines are fourth-order global polynomials fits (estimated separately on each side of the threshold). Observations to the right of the vertical dashed line are eligible for Fonden under the heavy rainfall criteria.

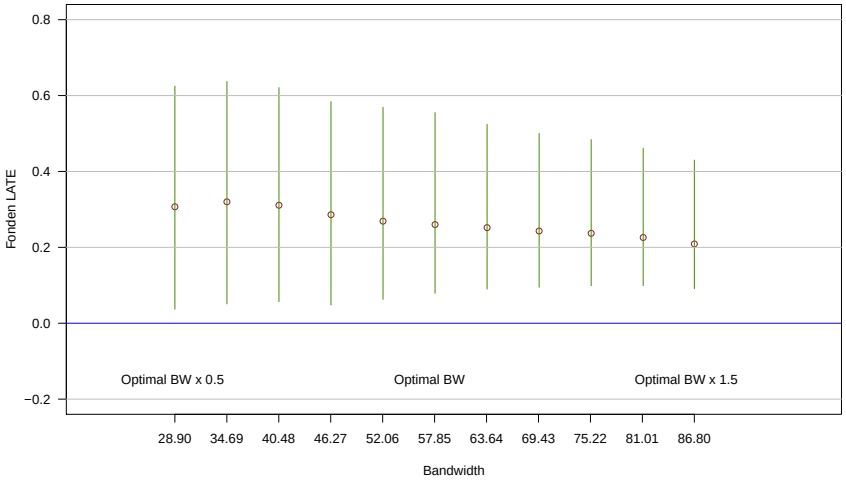


Figure A13: Fonden impact at various bandwidths

Note: Estimates of Fonden LATE at 10 evenly spaced bandwidths. The smallest bandwidth 28.6 mm is 50% smaller than the optimal h_{MSE} bandwidth, the largest 85.9 mm is 50% larger than the optimal h_{MSE}. The circles represent point estimates constructed using a triangular kernel, a local linear polynomial, and the bandwidth indicated in the axis. The error bars represent robust 95% confidence intervals.

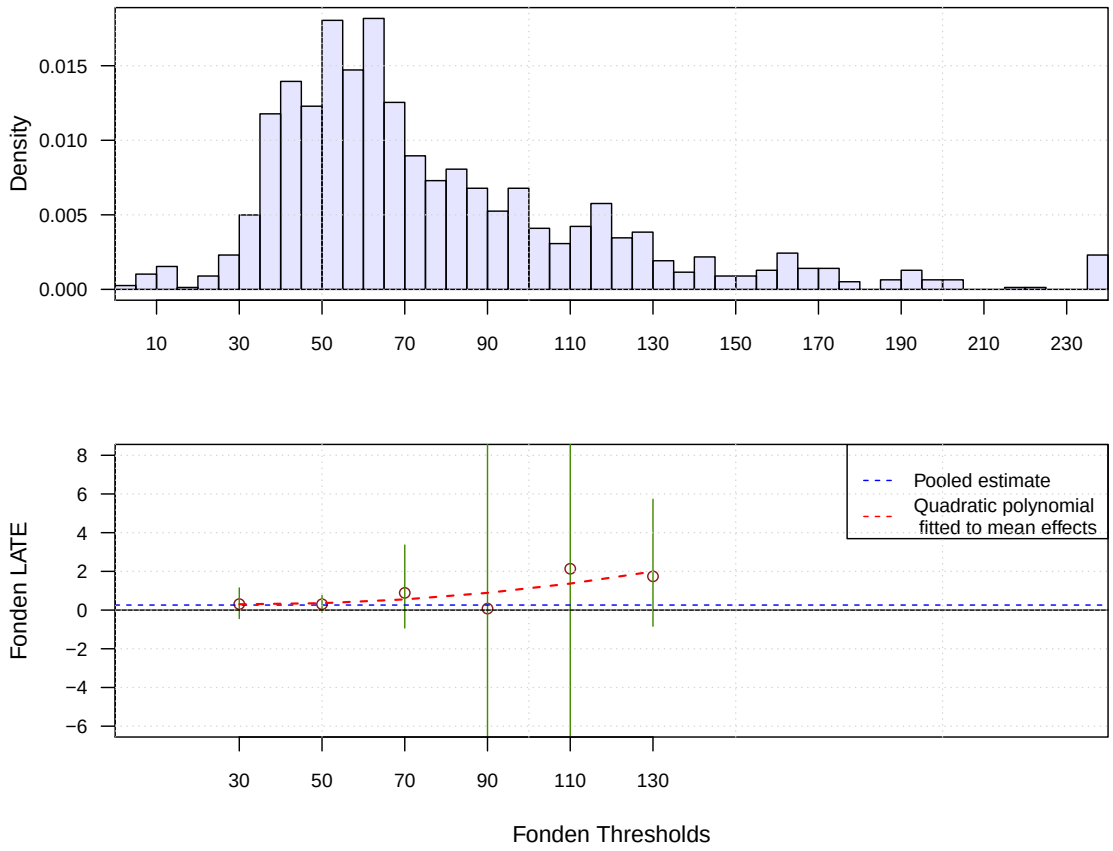


Figure A14: Fonden treatment effect curve

Note: The top panel plots the histogram of Fonden's Heavy rainfall thresholds. The bottom panel plots point estimates (circles) and 95% confidence intervals (green error bars) of Fonden's LATE at six threshold values. Each estimate of Fonden's LATE uses only the 400 treatment and control observations that are closest to each threshold. The bottom panel also plots Fonden's pooled LATE taken from table 2 column 1 (dashed blue lines), and a quadratic polynomial fit of Fonden's LATE at the six threshold values (red dashed lines).

References

Boudreau, Laura. 2015. "Discipline and disasters: The political economy of Mexicos Sovereign Disaster Risk Financing Program." FERDI. http://www.ferdi.fr/sites/www.ferdi.fr/files/publication/fichiers/b128_ferdi_l.boudreau.pdf.

Calonico, Sebastian, Matias D Cattaneo, and Rocio Titiunik. 2015. "Optimal data-driven regression discontinuity plots." *Journal of the American Statistical Association*, 110(512): 1753–1769.

Cattaneo, Matias D, Michael Jansson, and Xinwei Ma. 2018. "rddensity: Manipulation testing based on density discontinuity." *The Stata Journal*, 18(1): 234–261.

Cerulli, Giovanni, Yingying Dong, Arthur Lewbel, and Alexander Poulsen. 2017. "Testing Stability of Regression Discontinuity Models." In *M. D. Cattaneo and J. C. Escanciano (eds), Regression Discontinuity Designs: Theory and Applications (Advances in Econometrics, Volume 38, Chapter 4, pages 317-339)*. Emerald Group Publishing.

Conagua. 2015. "Daily Weather Station Historic Rainfall." Comision Nacional del Agua. ftp://200.4.8.36 (accessed March 3, 2015).

Conapo. 2005. "Infant mortality rate by municipality." Consejo Nacional de Poblacion. http://conapo.gob.mx/work/models/CONAPO/Resource/3cccfec6-7ab1-4a1b-9b28-5b731f3e51ef/tmi_mun_2005.html (accessed May 29, 2014).

DGIS. 2001. "Health Resources SINERHIAS 2001-2017." Secretaria de Salud, Direccion General de Informacion en Salud. http://www.dgis.salud.gob.mx/contenidos/basesdedatos/bdc_recursos_gobmx.html (accessed June 20, 2017).

Dong, Yingying, and Arthur Lewbel. 2015. "Identifying the effect of changing the policy threshold in regression discontinuity models." *Review of Economics and Statistics*, 97(5): 1081–1092.

Fonden. 2015. "Sistema Fonden en Linea." Secretaria de Hacienda y Credito Publico. http://www.proteccioncivil.gob.mx/en/ProteccionCivil/SistemaFondenenLinea (accessed March 20, 2015).

Henderson, J Vernon, Adam Storeygard, and David N Weil. 2011. "A bright idea for measuring economic growth." *The American Economic Review*, 101(3): 194.

INEGI. 2000. "Census 2000." Instituto Nacional de Estadistica y Geografia. https://www.inegi.org.mx/programas/ccpv/2000/ (accessed June 20, 2017).

INEGI. 2005. "Conteo 2005." Instituto Nacional de Estadistica y Geografia. https://www.inegi.org.mx/programas/ccpv/2005/ (accessed June 20, 2017).

INEGI. 2010. "Census 2010." Instituto Nacional de Estadistica y Geografia. https://www.

inegi.org.mx/programas/ccpv/2010/ (accessed June 20, 2017).

INEGI. 2013*a*. "Municipal boundaries." Instituto Nacional de Estadistica y Geografia. http://en.www.inegi.org.mx/app/biblioteca/ficha.html?upc=702825292829 (accessed May 29, 2014).

INEGI. 2013*b*. "State level GDP." Instituto Nacional de Estadistica y Geografia. https://www.inegi.org.mx/programas/pibent/2013/ (accessed May 29, 2014).

INEGI. 2014. "Public finances of municipalities." Instituto Nacional de Estadistica y Geografia. https://www.inegi.org.mx/programas/finanzas/ (accessed June 20, 2017).

USGS. 2003. "Global GIS : Global coverage DVD-ROM." United States Geological Survey: American Geological Institute. https://www.worldcat.org/title/global-gis-global-coverage-dvd-rom/oclc/68101735 (accessed October 10, 2017).

World Bank. 2010*a*. "GDP deflator." World Bank. https://data.worldbank.org/indicator/NY.GDP.DEFL.ZS (accessed July 24, 2016).

World Bank. 2010*b*. "PPP conversion factor, GDP (LCU per international $)." World Bank. https://data.worldbank.org/indicator/PA.NUS.PPP (accessed July 24, 2016).