Online Appendix for The Local Economic Impacts of Prisons^{*}

Janjala Chirakijja

Department of Economics, Northwestern University

JOB MARKET PAPER Please find the latest version HERE.

December 24, 2017

^{*}Email: janjalachirakijja2012@u.northwestern.edu. I am most grateful to my advisors Seema Jayachandran, Matthew Notowidigdo, Lori Beaman and Jonathan Guryan for their guidance and extraordinary support. This work has also benefited from stimulating discussions with Cynthia Kinnan, Jacopo Ponticelli, Christopher Udry, Treb Allen, David Berger, Yuta Takahashi, John Cawley, Lauren Velasco, and several seminar participants. I also thank Sachet Bangia for excellent research assistance and Caitlin Rowe and Andre Nickow for carefully reading my manuscript. All errors are my own.

1 Trimming and Matching to Improve Covariate Balance in the U.S. analysis

The U.S.-wide analysis, both at the census tract and the county level, is conducted with three samples — the full, trimmed, and matched sample. This section describes how I construct the trimmed and matched samples following Imbens and Rubin (2015). The approach applies to both the census tract and the county samples.

The first step is to estimate the propensity score for prison opening treatment. I follow Imbens and Rubin (2015) and select covariates and second-order terms from a rich set of pretreatment neighborhood characteristics to include in the propensity score using an iterative procedure.¹

This procedure selects 40 first-order and second-order terms at the census tract level²

²At the census tract level, these covariates are log median housing values, population density, proportion of population in correctional institutions, log total renter-occupied housing units, proportion of housing units with Hispanic householder, proportion of housing units with black householder, log median rental rates, proportion of over 16 in the civilian labor force, log mean household income, proportion of housing units that are single-unit detached, proportion with a college degree, proportion of housing units built within the last 5 years, proportion of housing units with three bedrooms, proportion of housing units with four bedrooms, proportion of housing units that are occupied, proportion of housing units built 6-20 years ago, proportion unemployed, proportion foreign born, proportion of families with children, proportion of population under age 18, proportion of housing units built 6-20 years ago interacted with proportion of population in correctional institutions, proportion of over 16 in the civilian labor force squared, proportion of families with children interacted with proportion of population in correctional institutions, proportion unemployed interacted with population density, log total renter-occupied housing units squared, proportion of families with children squared, proportion of housing units with Hispanic householder squared, population density interacted with proportion of population in correctional institutions, population density interacted with proportion of housing units that are occupied, proportion of housing units that are occupied squared, proportion of families with children interacted with proportion of housing units that are occupied, proportion of families with children interacted with proportion of housing units built within the last 5 years, proportion with a college degree interacted with proportion of housing units that are single-unit detached, proportion of housing units with

¹The procedure starts with a large set of potential covariates for including in the propensity score estimation. In the first step, we can decide to include some basic covariates on substantive grounds, which may include covariates that are *a priori* viewed as important for explaining the treatment and related to outcome measures. In this step, I select log median housing values in 1990 for automatic inclusion. The second step selects which of the remaining covariates to include in the specification of the propensity score. We start with a logit propensity score model with just an intercept and the covariates selected for automatic inclusion. Then we add each of the remaining covariates, one at a time, to the model, estimate the model, and calculate the likelihood ratio statistics assessing the null hypothesis that the newly included covariate has a zero coefficient. After repeating this for all potential covariates, we add the covariate with the largest likelihood ratio statistic to our specification. We start this process again with the remaining covariates and continue until all the likelihood ratio statistics are less than 1. In the third step, we decide which of the interactions and quadratic terms involving the first-order terms chosen in the second step to include in the specification of the propensity score. To do this, we repeat the iterative procedure described in the second step and keep including an additional second-order term until all of the remaining likelihood ratio statistics are less than 2.71.

and 32 first-order and second-order terms at the county level.³ I then estimate the propensity score for containing where a prison opened using these covariates in a logit regression.

The trimmed sample is then created by dropping observations with extreme values of the propensity score from the full sample, based on a systematic method proposed by Crump et al. (2009). Previous research has often used ad hoc criterion for trimming the sample. The Crump et al. (2009) approach looks for the subpopulation that allows for the most precise estimation of the average treatment effects within a selected subpopulation. If there is a value of the covariate such that there are few treated units relative to control units, then the variance for an estimator of the average treatment effect for that value of the covariate will be large. Hence, dropping units with such covariate values improves the asymptotic variance of the efficient estimator. This criterion leads to discarding observations with propensity scores outside an optimal threshold.⁴ The threshold is determined solely by the joint distribution of covariates and treatment status and does not depend on the distribution of outcomes, therefore no deliberate bias is introduced with regard to the treatment effect being analyzed.

 4 Crump et al. (2009) notes that the role of the propensity score in the selection rule is not imposed a priori, but emerges as a consequence of the criterion.

three bedrooms interacted with log total renter-occupied housing units, proportion of housing units built 6-20 years ago interacted with log median rental rates, proportion of housing units built 6-20 years ago interacted with log median housing values, proportion of housing units that are single-unit detached interacted with log mean household income, proportion of housing units with four bedrooms interacted with log median housing values, proportion of housing units squared, proportion of housing units built 6-20 years ago interacted with proportion of housing units that are occupied. All covariates are from the 1990 census.

³At the county level, the selected covariates are log median housing values, proportion of households female headed, proportion of housing units with four bedrooms, proportion of housing units with Hispanic householder, log total owner-occupied housing units, proportion of housing units with two bedrooms, proportion unemployed, proportion age 65 or older, population density, log mean household income, proportion without a high school diploma, proportion of housing units that are owner occupied, proportion of housing units consisting of three or four units, proportion of housing units consisting of two units, proportion of housing units built within the last 5 years, proportion of housing units with zero bedroom, log total owneroccupied housing units interacted with proportion of households female headed, log total owner-occupied housing units interacted with log median housing values, proportion of housing units consisting of two units interacted with proportion of housing units that are owner occupied, proportion of housing units that are owner occupied squared, proportion of housing units with four bedrooms interacted with log median housing values, proportion of housing units built within the last 5 years interacted with proportion of households female headed, proportion of households female headed squared, log mean household income interacted with proportion of housing units with four bedrooms, log mean household income interacted with log total owneroccupied housing units, proportion without a high school diploma interacted with proportion of households female headed, proportion of housing units that are owner occupied interacted with proportion of housing units with four bedrooms, proportion of housing units built within the last 5 years squared, proportion of housing units consisting of two units interacted with log median housing values, proportion of housing units consisting of three or four units interacted with proportion of housing units with Hispanic householder, proportion of housing units with Hispanic householder interacted with log median housing values, proportion of housing units consisting of three or four units interacted with proportion without a high school diploma. All covariates are from the 1990 census.

In the census tract analysis, using the optimal threshold, I exclude census tracts with predicted propensity score outside the interval [0.007, 0.993]. That is, trimming excludes 31 treated tracts and 47,461 control tracts. At the county level, I exclude 17 treated and 1,266 control counties with predicted propensity score outside the interval [0.036, 0.964]. Table A1 and A2 reports the summary statistics for the resulting trimmed census tract sample and trimmed county sample respectively.

The matched sample is constructed using propensity score matching. Specifically, I match each census tract (county) to the two census tracts (counties) with the closest propensity score, with replacement. These matched census tracts (counties) form the control group of my matched sample. The summary statistics for the matched census tract sample and county sample are presented in the main text.

2 U.S. Estimates Based on Alternative Strategies and Samples

Table A3 reports the estimated impact of prison openings on median housing values, median rents, and average household income using alternative strategies. Each entry gives the estimated impact on the outcome presented in each row. Column (1) and (2) use the full sample of census tracts; Columns (3) and (4) use the trimmed sample and Columns (5) and (6) use the propensity score matched sample. The odd columns report the estimates based on the following fixed effects specification.

$$\Delta y_{i2000} = \alpha T_{i2000} + X'_{i1990}\beta + \epsilon_{i2000} \tag{A1}$$

 Δy_{i2000} is the change in some outcome over the period 1990-2000 in community *i*. The indicator variable T_{i2000} takes the value 1 for census tracts where a prison opened between 1990 and 2000. Vector X_{i1990} includes a rich set of pre-treatment housing and demographic controls. ϵ is the unobservable components of the outcome.

The even columns report the estimates based on the Parametric Reweighting approach following Busso, Gregory, and Kline (2013). Specifically, I estimate interacted regressions of the following form to allow for flexible patterns of treatment effect heterogeneity.

$$\Delta y_{i2000} = \mu^1 T_{i2000} + (1 - T_{i2000}) \times X'_{i1990}\beta + \epsilon_{i2000}, \tag{A2}$$

where $\mu^1 \equiv E[\Delta y_{i2000}|T_{i2000} = 1]$. This specification models the mean change in outcomes among the control tracts as a linear function of X_{i1990} , but is agnostic regarding the conditional expectation function among the treated tracts. An estimate of the average treatment effect on treated tracts (\widehat{ATT}) can be formed as

$$\widehat{ATT} = \hat{\mu}^1 - \frac{1}{N_1} \sum_i T_{i2000} X'_{i1990} \hat{\beta}.$$
(A3)

where N_1 is the number of treated tracts. The reported standard errors in Table A3 are clustered by state-specific urban/rural area.

Overall, the results are consistent with my baseline results although the Parametric Reweighting estimates are imprecise. Fixed effects estimate and lagged dependent variable estimate (i.e., my baseline estimate) have a useful bracketing property so finding similar results adds credibility to my results.

Table A4 reports estimates from a different robustness check, using alternative estimating samples. In this analysis, I use my baseline specification as described in the main text. That is, I estimate:

$$y_{i2000} = \alpha T_{i2000} + X'_{i1990}\beta + \epsilon_{i2000}, \tag{A4}$$

where y_{i2000} is the outcome in 2000 in community *i*, and T_{i2000} and X_{i1990} are as defined above except that vector X_{i1990} in this specification also includes the lagged dependent variable.

However, instead of using all census tracts in the United States as in my baseline estimation, I restrict the control group of the full sample to include only census tracts in counties that experienced prison opening during 1970-2000. The trimmed and matched samples are constructed as before by dropping some tracts from the full sample. Therefore, the trimmed and matched samples are also restricted to tracts in counties that experienced prison openings two decades ago or during the 1990s. The motivation for this exercise is that communities that experienced prison openings in the past are likely to share unobserved characteristics and trends with communities that experienced prison openings during the decade of interest.

All columns in Table A4 control for the 1990 dependent variable, the complete set of housing and demographic characteristics from 1990 and state-urban area fixed effects. The third column also adds tract-level housing and demographic characteristics from 1980.⁵ The estimates are consistent with my baseline results.

3 Constructing Control Group for the Texas Quasi-Experimental Analysis

This section describes how I construct the analysis sample of census tracts, referred to as the "site-selection" sample, for the Texas quasi-experimental analysis. The approach selects tracts in the rejected counties with a sufficient likelihood of containing a prison site (had their counties won) to form my control group.

First, I estimate the propensity score for containing the prison site using all census tracts in the winning counties. In this step, I again follow Imbens and Rubin (2015) and select first and second order covariates to include in the propensity score using an iterative procedure.⁶ Second, I use this propensity score specification to predict the propensity score for all census tracts in the rejected counties. Finally, census tracts in the rejected counties with a predicted propensity score between the minimum and the maximum of the predicted propensity score among treated tracts (i.e., common support) are selected to be included in the control group of my analysis sample.⁷

I turn next to discuss a robustness exercise which shows that my estimates of the impacts of prisons are robust to an alternative approach to constructing the control group. This approach uses propensity score matching to construct the control group instead of using

⁷That is, census tracts in the rejected counties with a predicted propensity score in the interval [.004,0.927].

 $^{^{5}}$ As noted in the data section, nonurban areas were not tracted in 1980 so data are not available for nonurban Census tracts in 1980. To keep all observations, the specifications in my even columns also include dummy variables for missing value of 1980 controls, one for each of the 1980 characteristics.

⁶These covariates are log median housing values, population density, log total renter-occupied housing units, proportion unemployed, proportion below the poverty line, proportion of housing units with zero bedroom, proportion with a college degree, proportion of housing units with two bedroom, proportion over age 65, proportion Hispanic, proportion of housing units that are occupied, proportion black, proportion black interacted with proportion over age 65, proportion black squared, proportion of housing units that are single-unit detached interacted with log total renter-occupied housing units, proportion of housing units that are single-unit detached interacted with proportion of housing units that are occupied, proportion Hispanic interacted with proportion over age 65, proportion black interacted with proportion Hispanic interacted with proportion over age 65, proportion black interacted with proportion with a college degree, proportion of housing units that are single-unit detached interacted with proportion of housing units with two bedroom, proportion of housing units that are single-unit detached squared, proportion of housing units with two bedroom, proportion of housing units that are single-unit detached squared, proportion of housing units with one bedroom interacted with proportion Hispanic, proportion of housing units that are single-unit detached squared, proportion of housing units with one bedroom interacted with proportion of housing units with one bedroom interacted with proportion Hispanic, proportion of housing units that are single-unit detached interacted with population density, and proportion of housing units that are single-unit detached interacted with proportion with a college degree. All covariates are from the 1990 census.

census tracts in the rejected counties with a predicted propensity score between the minimum and the maximum of the predicted propensity score among treated tracts. Specifically, I match each treated census tract in the winning counties to the four census tracts in the rejected counties with the closest propensity score, with replacement. These matched census tracts form the control group of this alternative sample. Table A5 reports the estimates of the impacts of prisons on my key outcomes using the full sample and the alternative "matched site-selection" sample of Texan census tracts. The results are similar to my baseline results.

4 Texas Results for Other Outcomes

This section presents the rest of the results for Texas based on my quasi-experimental approach that are omitted from the main text for brevity.

Table A6 provides estimates of the impact of prison openings on average household income at the census tract level by when householders moved into the housing units, including in the past 10 years, 10-20 years ago and 20 or more years ago. The estimate in Column (2) and (4) indicate that compared to neighborhoods with similar characteristics, neighborhoods that experience prison openings see 6-7 percent decreases in the average income of households who recently moved in in the past 10 years and who have been in the housing units more than 20 years ago. The estimate for households who have been residing in the neighborhoods 10-20 years is not statistically different from zero. Although movers may include existing neighborhood residents who relocate within the same neighborhood, these results suggest that the overall decreases in average household income (discussed in the main text and reported again for reference in Column (1) of Table A6), are driven by recent in-migration of households with lower income and out-migration of long-time residents with higher income following the opening of a prison.

I turn next to describe the figures and tables from the county-level analysis for Texas. Figure A1 to A3 plot the coefficients associated with the set of treatment indicators, one for each of the 5 leads and 5 lags and the contemporaneous year of prison openings based on my difference-in-differences analyses. Overall, the results are consistent with those based on the U.S.-wide analyses. There is no evidence of a trend in employment, average wages, or number of establishments prior to the year prisons were opened in the treated counties relative to the other counties, lending credibility to my identification strategy. Following the openings of prisons, treated counties saw large and persistent gains in government employment but little change in employment and number of establishments in the private sector. There is also evidence that prison buildings create a temporary boost to the local construction industry raising employment and average wages 1-2 years ahead of when prisons were opened or during the construction period.

Finally, Table A7 reports the county-level estimates for the impact of prisons on housing values, demographic characteristics and local government finances. As noted in the main text, power is lower with this quasi-experimental approach compared to the U.S. wide approach. For the outcomes presented in this table, power represents an important limitation and relatively small effect sizes cannot be ruled out.⁸

⁸For example, the minimum detectable effect sizes (MDEs) for housing values and household income based on column (2) and (3) lie in the range of 6-10 percent.

References

- Busso, M., J. Gregory, and P. Kline (2013). "Assessing the Incidence and Efficiency of a Prominent Place Based Policy". *American Economic Review* 103 (2), pp. 897–947.
- Crump, R. K., V. J. Hotz, G. W. Imbens, and O. A. Mitnik (2009). "Dealing with limited overlap in estimation of average treatment effects". *Biometrika* 96 (1), pp. 187–199.
- Imbens, G. W. and D. B. Rubin (2015). *Causal Inference for Statistics, Social, and Biomedical Sciences: An Introduction.* Cambridge University Press.

Figure A1: Impact of prisons on local employment in the state and local government sector, Texas counties



Notes: This figure plots results from the difference-in-differences analysis allowing for 5 leads and 5 lags for log of employment in the state and local government sector. The sample consists of Texas counties in the site selection sample. County-year data from 1985 to 2005 for the reported outcome come from the Quarterly Census of Employment and Wages (QCEW). The bars report 95 percent confidence intervals calculated using standard errors clustered at the county level.

Figure A2: Impact of prisons on local employment, wages, and number of establishments in the private sector, Texas counties



Notes: This figure plots results from the difference-in-differences analysis allowing for 5 leads and 5 lags for log of employment, log of average weekly wages, and log of number of establishments in the private sector. The sample consists of Texas counties in the site selection sample. County-year data from 1985 to 2005 for the reported outcomes come from the Quarterly Census of Employment and Wages (QCEW). The bars report 95 percent confidence intervals calculated using standard errors clustered at the county level.



Figure A3: Temporary boost to local construction industry, Texas counties

Notes: This figure plots results from the difference-in-differences analysis allowing for 5 leads and 5 lags for log of employment and log of average weekly wages in the construction industry. The sample consists of Texas counties in the site selection sample. County-year data from 1985 to 2005 for the reported outcomes come from the Quarterly Census of Employment and Wages (QCEW). The bars report 95 percent confidence intervals calculated using standard errors clustered at the county level.

Table A1: Pre-treatment mean characteristics of census tracts in the U.S. in the full and trimmed sample

	Full		Trimmed		Normalize	ed Diff. & t-test
	(1)	(2)	(3)	(4)	(5)	(6)
	Treated	Control	Treated	Control	Full	Trimmed
Housing Characteristics						
Median House value (\$)	50,702.29	89,650.48	47,949.00	48,850.25	0.94^{***}	0.04
Monthly rent (\$)	315.07	462.76	309.37	308.20	1.06***	-0.01
Proportion owner occupied	0.62	0.60	0.62	0.64	-0.14*	0.19^{***}
Proportion occupied	0.86	0.91	0.86	0.87	0.48^{***}	0.10
Proportion 0-1 bedrooms	0.10	0.14	0.11	0.10	0.36^{***}	-0.13**
Proportion 2-3 bedrooms	0.77	0.70	0.78	0.78	-0.67***	-0.03
Proportion 4+ bedrooms	0.12	0.16	0.11	0.13	0.38^{***}	0.20^{***}
Proportion consisting of 1 unit	0.70	0.67	0.69	0.71	-0.19^{**}	0.17^{***}
Proportion consisting of 2 unit	0.03	0.05	0.03	0.03	0.38^{***}	0.01
Proportion consisting of 3-4 unit	0.02	0.05	0.02	0.02	0.50^{***}	-0.00
Proportion consisting of $5+$ unit	0.05	0.15	0.05	0.04	0.69^{***}	-0.09
Proportion mobile home	0.19	0.08	0.20	0.18	-0.90***	-0.10
Proportion complete plumbing	0.85	0.90	0.84	0.85	0.56^{***}	0.10
Proportion built last 5 yrs	0.10	0.11	0.10	0.10	0.15^{*}	-0.05
Proportion built within 6-20 yrs	0.36	0.31	0.37	0.37	-0.30***	-0.00
Proportion built within 20-40 yrs	0.28	0.30	0.29	0.27	0.14^{*}	-0.17***
Proportion built >40 yrs	0.25	0.27	0.24	0.27	0.07	0.15^{**}
Demographic Characteristics						
Household annual income (\$)	$28,\!483.77$	36,999.45	28,030.51	27,859.87	0.75^{***}	-0.02
Population density (/sq.km)	94.14	$1,\!650.90$	67.45	71.63	0.67^{***}	0.03
Proportion no high school diploma	0.35	0.25	0.37	0.35	-0.76***	-0.15**
Proportion with college	0.10	0.19	0.10	0.10	0.85^{***}	0.08
Proportion under 18	0.27	0.26	0.27	0.27	-0.25***	0.06
Proportion 65+	0.14	0.13	0.14	0.14	-0.16**	0.08
Proportion houshold black	0.13	0.11	0.14	0.11	-0.09	-0.17^{***}
Proportion houshold hispanic	0.07	0.06	0.08	0.05	-0.05	-0.21***
Proportion below poverty line	0.19	0.13	0.20	0.18	-0.61^{***}	-0.18***
Proportion public assistance	0.10	0.08	0.10	0.10	-0.40***	-0.11^{*}
Proportion foreign born	0.03	0.07	0.04	0.03	0.40^{***}	-0.15**
Proportion female headed	0.21	0.22	0.22	0.20	0.04	-0.18***
Total employment $(16 + \text{ persons})$	$1,\!695.52$	1,829.21	1,667.51	$1,\!648.29$	0.16^{**}	-0.03
Unemployment rate	0.08	0.07	0.08	0.08	-0.28***	-0.03
Proportion in correctional inst.	0.03	0.00	0.04	0.01	-0.33***	-0.26***
Observations	255	56,766	224	9,305		

Notes: Column (1) reports mean characteristics for treated tracts in the full sample where prisons were opened during the 1990s. Column (2) describes characteristics for control tracts in the full sample. Column (3) and (4) describe characteristics for treated and control tracts in the trimmed sample respectively. The numbers reported in Column (5) to (6) are normalized differences between treated and control tracts in each of my samples, defined as the difference in means scaled by the square root of the average of the two within-group variances. The asterisks report the significance levels of the t-tests of mean equality between treated and control tracts in the corresponding sample, with one, two and three asterisks indicating a p-value less than 0.1, 0.05 and 0.01 respectively.

Table A2: Pre-treatment mean characteristics of counties in the U.S. in the full and trimmed sample

	Fi	ıll	Trin	nmed	Normalized Diff. & t-test	
	(1) Treated	(2) Control	(3) Treated	(4) Control	(5) Full	(6) Trimmed
Housing Characteristics						
Median House value (\$)	44,883.56	54,138.30	43,248.51	46,879.70	0.34***	0.25^{***}
Monthly rent (\$)	298.65	322.33	293.24	304.25	0.29***	0.18^{**}
Proportion owner occupied	0.62	0.62	0.61	0.62	0.03	0.14^{*}
Proportion occupied	0.85	0.85	0.85	0.86	-0.01	0.11
Proportion 0-1 bedrooms	0.11	0.11	0.11	0.10	-0.03	-0.12^{*}
Proportion 2-3 bedrooms	0.77	0.74	0.77	0.77	-0.39***	-0.11
Proportion 4+ bedrooms	0.12	0.15	0.12	0.13	0.51^{***}	0.27^{***}
Proportion consisting of 1 unit	0.72	0.73	0.72	0.72	0.12^{*}	0.03
Proportion consisting of 2 unit	0.03	0.03	0.03	0.03	-0.05	-0.03
Proportion consisting of 3-4 unit	0.03	0.03	0.02	0.03	0.17^{**}	0.13^{*}
Proportion consisting of $5+$ unit	0.05	0.06	0.05	0.05	0.14^{*}	0.07
Proportion mobile home	0.16	0.14	0.16	0.16	-0.23***	-0.08
Proportion complete plumbing	0.84	0.84	0.83	0.84	0.03	0.14^{*}
Proportion built last 5 yrs	0.09	0.10	0.09	0.09	0.07	0.06
Proportion built within 6-20 yrs	0.34	0.34	0.34	0.35	-0.00	0.07
Proportion built within 20-40 yrs	0.28	0.26	0.28	0.27	-0.24***	-0.13^{*}
Proportion built >40 yrs	0.28	0.30	0.29	0.29	0.09	0.00
Demographic Characteristics						
Household annual income (\$)	$27,\!553.34$	$29,\!611.26$	$27,\!095.46$	$28,\!104.22$	0.33^{***}	0.22^{***}
Population density (/sq.km)	56.97	86.06	53.09	45.13	0.06	-0.03
Proportion no high school diploma	0.35	0.30	0.36	0.33	-0.44***	-0.28***
Proportion with college	0.12	0.14	0.11	0.12	0.31^{***}	0.18^{**}
Proportion under 18	0.28	0.27	0.28	0.27	-0.22***	-0.17**
Proportion 65+	0.14	0.13	0.14	0.13	-0.04	-0.16**
Proportion houshold head black	0.12	0.07	0.13	0.10	-0.33***	-0.17^{**}
Proportion houshold head hispanic	0.07	0.03	0.07	0.04	-0.29***	-0.27***
Proportion below poverty line	0.20	0.16	0.21	0.18	-0.45^{***}	-0.30***
Proportion public assistance	0.10	0.08	0.11	0.09	-0.45^{***}	-0.28***
Proportion foreign born	0.02	0.02	0.02	0.02	-0.07	-0.15^{**}
Proportion female headed	0.21	0.18	0.22	0.20	-0.46^{***}	-0.23***
Unemployment rate	0.08	0.07	0.08	0.07	-0.36***	-0.17^{**}
Proportion in correctional inst.	0.01	0.01	0.01	0.01	-0.03	0.10
Observations	219	2,877	202	$1,\!611$		

Notes: Column (1) reports mean characteristics for treated counties in the full sample where prisons were opened during the 1990s. Column (2) describes characteristics for control counties in the full sample. Column (3) and (4) describe characteristics for treated and control counties in the trimmed sample respectively. The numbers reported in Column (5) to (6) are normalized differences between treated and control counties in each of my samples, defined as the difference in means scaled by the square root of the average of the two within-group variances. The asterisks report the significance levels of the t-tests of mean equality between treated and control counties in the corresponding sample, with one, two and three asterisks indicating a p-value less than 0.1, 0.05 and 0.01 respectively.

	Full Sa	Full Sample		l Sample	Matched Sample		
	(1)FE	(2) PW	(3)FE	$\begin{pmatrix} 4 \\ PW \end{pmatrix}$	(5)FE	(6) PW	
Log(House value)	-0.034^{**} (0.016)	-0.034 (0.036)	-0.035^{**} (0.015)	-0.035 (0.037)	$\begin{array}{c} -0.041^{***} \\ (0.012) \end{array}$	-0.050 (0.038)	
Log(Rent)	$0.005 \\ (0.012)$	$0.004 \\ (0.016)$	$0.009 \\ (0.011)$	$0.006 \\ (0.017)$	-0.001 (0.015)	-0.005 (0.018)	
Log(HH Income)	-0.023^{***} (0.009)	-0.023^{*} (0.013)	-0.028^{***} (0.009)	-0.029^{**} (0.013)	-0.018^{**} (0.009)	-0.017 (0.013)	
Observations 1990 characteristics	57,023 Yes	57,023 Yes	9,529 Yes	9,529 Yes	744 Yes	744 Yes	

Table A3: U.S. tract-level robustness: Fixed effects and Parametric Reweighting approaches

Notes: Each entry reports the estimate of the impact of prison openings on the outcome presented in each row. Odd columns report the estimates based on the fixed effects specification. Even columns report the estimates based on the Parametric Reweighting approach following Busso, Gregory, and Kline (2013). Reported standard errors are clustered by state-urban/rural area. Asterisks denote significance: * p < .10, ** p < .05, *** p < .01.

Table A4: U.S. tract-level robustness: Control tracts in counties with prison openings during 1970-2000

	Full Sample		Trimmed	l Sample	Matched Sample	
	(1)	(2)	(3)	(4)	(5)	(6)
Log(House value)	-0.043^{***} (0.011)	-0.033^{***} (0.011)	-0.032^{***} (0.012)	-0.037^{***} (0.013)	-0.026^{**} (0.013)	-0.023^{*} (0.013)
Log(Rent)	-0.021^{**} (0.011)	-0.011 (0.011)	$0.001 \\ (0.010)$	-0.003 (0.011)	-0.002 (0.011)	$0.002 \\ (0.011)$
Log(HH Income)	-0.037^{***} (0.009)	-0.023^{**} (0.010)	-0.030*** (0.008)	-0.027^{***} (0.008)	-0.020** (0.009)	-0.014 (0.009)
Observations	17,950	$17,\!950$	$1,\!697$	$1,\!697$	685	685
1990 dependent variable	Yes	Yes	Yes	Yes	Yes	Yes
1990 characteristics	Yes	Yes	Yes	Yes	Yes	Yes
1980 characteristics	No	Yes	No	Yes	No	Yes
State-Urban FEs	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Each entry reports the estimate of the impact of prison openings on the outcome presented in each row. Heteroskedasticity-robust standard errors are reported in parentheses. Asterisks denote significance: * p < .10, ** p < .05, *** p < .01. In all three samples, the control group is restricted to census tracts in counties that experienced prison openings during 1970-2000.

	Full	Site Selecti	on (Matched)
	(1)	(2)	(3)
Panel A: Housing Impacts			
Log(House value)	-0.007 (0.030)	$0.008 \\ (0.035)$	$0.039 \\ (0.040)$
Log(Rent)	$\begin{array}{c} 0.005 \\ (0.034) \end{array}$	$0.010 \\ (0.029)$	$0.011 \\ (0.037)$
Panel B: Income & Demographic Impacts			
Log(HH Income)	-0.030^{*} (0.017)	-0.055^{***} (0.019)	-0.040^{*} (0.021)
Log(HH Wage & Salary)	-0.043^{**} (0.018)	-0.070^{***} (0.021)	-0.043^{*} (0.022)
Prop. Black	-0.003 (0.004)	$0.005 \\ (0.006)$	$0.005 \\ (0.007)$
Prop. Hispanic	0.010^{*} (0.006)	0.011^{*} (0.006)	$0.005 \\ (0.007)$
Prop. Owner Occupied	-0.011 (0.008)	-0.024^{***} (0.007)	-0.023^{***} (0.008)
Panel C: Population & Employment Impacts			
Log(Population in HH)	$0.002 \\ (0.024)$	-0.032 (0.027)	$0.011 \\ (0.029)$
Log(Employment)	$0.030 \\ (0.026)$	-0.011 (0.031)	$0.027 \\ (0.036)$
Unemployment rate	-0.009^{**} (0.004)	-0.007 (0.004)	-0.006 (0.005)
Observations	3,156	138	138
1990 dependent variable	Yes	Yes	Yes
1990 characteristics	Yes	Yes	Yes
1980 characteristics	No	No	Yes

Table A5: Impact of prisons on housing, demographic and employment outcomes, Texas census tracts (Propensity score matching)

Notes: Each entry reports the estimate of the impact of prison openings on the outcome presented in each row. Site selection sample is constructed using propensity score matching, where each of the treated tracts is matched to the four tracts in the rejected counties with the closest propensity score. All columns control for the full set of 1990 housing and demographic characteristics. Column (3) adds the following 1980 characteristics: log median housing values, log median rental rates, proportion of housing units that are owner occupied, population density, proportion of housing units with Hispanic householder, proportion of housing units with black householder, proportion with no high school diploma, log average household income, log average household wages and salary, proportion unemployed, proportion under age 18, and proportion over age 65. Heteroskedasticity-robust standard errors are reported in parentheses. Asterisks denote significance: * p < .10, ** p < .05, *** p < .01.

		Log(Average Income)					
	(1) All	(2) Past 10 yrs	(3) 10-20 yrs	(4) $20+ yrs$	(5) Past 10 /10+		
Prison opening	-0.0608*** [0.0205]	-0.0675*** [0.0243]	0.00467 [0.0384]	-0.0719^{*} [0.0410]	-0.0355 $[0.0366]$		
Observations	233	233	233	233	233		
R-squared	0.848	0.818	0.648	0.589	0.294		
Control mean in 2000 (Level)	47,029	46,775	50,908	$42,\!946$	1.01		
1990 Log(Average Income)	Yes	Yes	Yes	Yes	Yes		
1990 characteristics	Yes	Yes	Yes	Yes	Yes		
State-Urban FEs	Yes	Yes	Yes	Yes	Yes		

Table A6: Impact of prisons on average household income by year householders moved in, Texas census tracts

Notes: Column (1) to (4) report the estimates of the impact of prison openings on average household income by when householders moved into the housing units, including in the past 10 years, 10-20 years ago and 20 or more years ago, respectively. Column (5) reports the estimates the impact of prison openings on the ratio of average income among households who moved in in the past 10 years to that among households who moved in more than 10 years ago. All estimates in this table are based on the site selection sample of census tracts. Heteroskedasticity-robust standard errors are reported in parentheses. Asterisks denote significance: * p < .10, ** p < .05, *** p < .01.

	Full	Site Se	election
	(1)	(2)	(3)
Panel A: Housing Impacts			
Log(House value)	$0.006 \\ (0.009)$	$0.033 \\ (0.024)$	$0.016 \\ (0.033)$
Log(Rent)	0.010^{*} (0.005)	0.013 (0.022)	-0.005 (0.025)
Panel B: Demographic Impacts			
Log(HH Income)	$0.003 \\ (0.005)$	-0.003 (0.020)	-0.004 (0.021)
Prop. black HHs	-0.003^{***} (0.001)	$0.002 \\ (0.002)$	$0.001 \\ (0.003)$
Prop. Hispanic HHs	0.002^{*} (0.001)	$0.003 \\ (0.005)$	$0.007 \\ (0.005)$
Panel C: Local Government Finance Impacts			
Log(Revenue per capita)	$0.007 \\ (0.012)$	-0.032 (0.051)	-0.016 (0.063)
Log(Intergov. Revenue per capita)	$0.021 \\ (0.014)$	-0.064 (0.048)	-0.096 (0.061)
Log(Expenditure per capita)	0.003 (0.012)	-0.004 (0.054)	$0.002 \\ (0.066)$
Observations	$3,\!093$	79	79
1990 dependent variable	Yes	Yes	Yes
1990 characteristics	Yes	Yes	Yes Voc

Table A7: Impact of prisons on housing and demographic characteristics, and local government finances, Texas counties

Notes: Each entry reports the estimate of the impact of prison openings on the outcome presented in each row. In Panel C, variables described as per capita refer to the line item divided by the county population that resides in households, which excludes incarcerated individuals. All columns control for the full set of 1990 housing and demographic characteristics. Column (3) adds the following 1980 characteristics: log median monthly rents, share of households with head black, share of households with head Hispanic, proportion with no high school diploma, log average household income, log employed individuals over 16 and unemployment rate. Heteroskedasticity-robust standard errors are reported in parentheses. Asterisks denote significance: * p < .10, ** p < .05, *** p < .01.