Does the Gender Wage Gap Influence Intimate Partner Violence in Brazil?
Evidence from Administrative Health Data

Elizaveta Perova, Sarah Reynolds, & Ian Schmutte

Online Appendix

APPENDIX

Informal sector comparison

The analysis uses the RAIS data to generate the gender wage ratio. This matched employer-employee dataset contains all labor market contracts in the formal sector in Brazil. However, in Brazil, 20 percent of laborers work in the informal sector.[xv] In this section, we use the nationally representative household survey, PNAD, to compare the gender wage ratios in the formal and informal sectors and to discuss how the differences might affect the main results.

The 2011–15 editions of PNAD are used with a few sample restrictions. First, the focus is on individuals aged 15–49, which is the age-group of interest in the main analysis. Second, the following categories of workers are excluded from the validation exercise: the self-employed, farm laborers, and domestic workers, which is similar to the sample restriction applied by Gerard et al. (2018) in a similar validation exercise. The analysis classifies workers as informal if they report that they have worked in the previous week, but do not have formal labor contracts.

The leave-out gender wage ratio constructed using the PNAD data has a few differences with the measures constructed based on RAIS. First, PNAD data are not as granular as the RAIS data. Measures are therefore calculated at the state level instead of the municipal level. Second, the industry variable reflects the broad industry definitions of PNAD, which consist of 12 categories, such as agriculture and construction. The analysis continues to use the 2011 baseline shares in each industry as the fixed proportions.

The distributions of the gender wage ratio in the formal and informal sectors overlap closely if one considers the raw wage ratios (Figure A1). However, the leave-out wage ratio is higher in the informal sector than in the formal sector, suggesting that there is less gender inequality in the informal sector if one is considering a measure that reflects wage changes in other states. Using the raw measure, there are more positive changes between 2011 and 2015 in the informal sector than in the formal sector. Yet, the same conclusion does not hold if one looks at the leave-out wage ratio. A possible explanation is that most of the improvement in the gender wage ratio was driven by the sorting of women into more profitable industries, which is the component that is kept fixed in computing the leave-out measure. This suggests that omitting the information on wages in the informal sector is not influencing trends of the leave-out measure, however, greater equality in informal wages (compared to formal wages) suggests we are under-stating equality, which may dampen impacts of movements in the formal sector in our analysis. Thus, we consider our estimates lower bounds.

The analysis also considered whether men or women are better represented in the informal sector. It accomplished this by regressing on sex an indicator of activity in a formal sector job. The analysis controlled for tenure, age, age squared, race, and educational attainment. It also includes year and state fixed effects. It found that women are 1 percent more likely to hold a formal sector job (Table A1). Even though the null hypothesis may be rejected that women and men do not sort selectively into the informal sector, the magnitude of this difference is small.

Given that a difference in differences specification is being used to measure the effects of the gender wage ratio on violence against women, only time-varying changes in wages in the informal sector that are not captured by aggregate time trends would lead to a potential bias in the estimates. Thus, the analysis examined if women's wages in the informal sector behave differently than men's wages in the same sector over time. It found that wages are lower in the informal sector among both women and men, slightly more so for men (Table A2). However, it does not appear that women's wages in the informal sector behave differently over time than the corresponding men's wages.

Robustness exercises

We examined robustness of results to using homicide rates per 100,000 women ages 15–

49 as outcome. In all cases, the rates and counts were also Winsorized. We also check robustness to a count data model that can more formally account for the presence of observations where the count is zero. Since there interest in our paper is on measuring the marginal effect of the gender wage ratio on the conditional mean outcome, we use a Poisson model. The Poisson model is preferred to other count data models because it is consistent under conditions similar to OLS, robust to various forms of misspecification, and it is easy to adjust inference for heteroskedasticity (Wooldridge 1999; Correia, Guimarães, and Zylkin 2019; 2020). The results in Table A3 show that the sign and magnitude of the coefficients on the leave-out wage ratio in the Poisson models for each outcome were consistent with the main results and generally more precisely estimated.[xvi]

Without population weights, but with controls, the coefficient on the leave-out wage ratio for the homicide outcome is not significant but remains negative and of similar magnitude (Table A4). This confirms a few large municipalities are not driving the results.

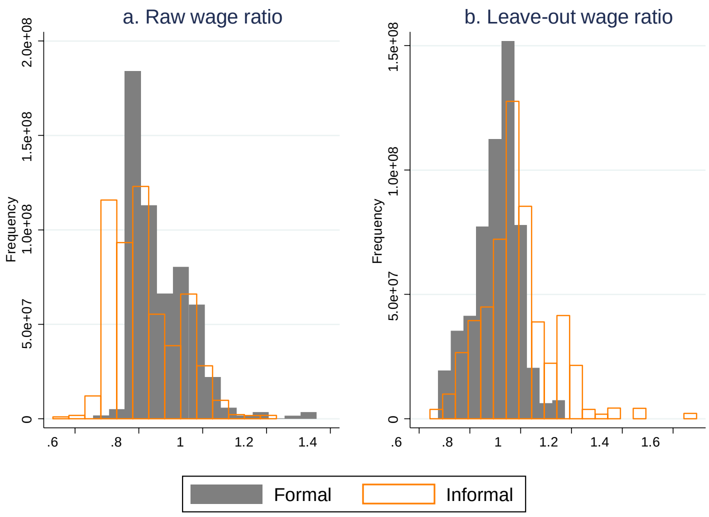
To further test that our leave-out wage ratio could still be endogenous to IPV, following Aizer (2010) we create a Bartik-style instrument that does not use information on wages at all. Instead, we predict employment for each gender based on state-wide sectoral growth (holding out the focal municipality) allocated according to the 2011 distribution of municipal employment across sectors. We then use the predicted male and female employment as instruments for the raw wage ratio. We find the results are broadly consistent, with a somewhat stronger support for an effect of increased gender wage parity in reducing homicides (Table A5). Curiously, the IV model shows a statistically significant positive effect on estimated IPV reports when controls are included in the model. This may reflect a local average treatment effect centered on particular municipalities (perhaps those with women's police stations).

Another robustness check increased the sample size to include the top two quintiles of the population rather than only the top quintile. The coefficients remained negative and became significant in the case of homicides (Table A6).

Some of the control variables—those indicative of women's empowerment—may be jointly determined or potential mediating variables, rather than true confounders. Thus, the analysis contrasts results with and without these sets of controls and finds consistent estimates across specifications (Table A7).

Including state trends, an interaction between state indicator variables and a continuous year variable resulted in even stronger magnitude and significance in the homicide outcome (Table A8). This result suggests that in states where female homicide is increasing, the female-male wage ratio is falling, and hence the leave-out wage ratio is also falling. This could reflect an unmodeled relationship between shifts in the industrial composition of state economies and homicide rates, which are associated with economic conditions and opportunities. By controlling for state-specific trends, we identify the effect of a rising female-male wage ratio on female homicides using variation in the wage ratio relative to the trend. The robustness results show that in municipalities where gender wage ratios are high relative to the trend homicide rates are low relative to the trend.

Figure A1: The gender wage ratio in formal and informal sectors



Note: Wage ratios are calculated by state-year for 2011–15 using PNAD data. Frequencies are adjusted by population weight.

Table A2: Changes in income from the informal sector over time, by sex

	(1) Log wage	(2) Log wage	(3) Log earnings	(4) Log earnings
Informal sector (1/0)	-0.318***	-0.312***	-0.316***	-0.309***
	(0.033)	(0.033)	(0.034)	(0.033)
Informal x (Year = 2012)	-0.002	-0.008	0.000	-0.008
	(0.009)	(0.012)	(0.008)	(0.012)
Informal x (Year = 2013)	0.015*	0.004	0.017*	0.005
	(0.006)	(0.009)	(0.007)	(0.009)
Informal x (Year = 2014)	0.011	0.008	0.012	0.008
	(0.009)	(0.011)	(0.009)	(0.011)
Informal x (Year = 2015)	-0.013	-0.027*	-0.008	-0.020
	(0.010)	(0.012)	(0.011)	(0.012)
Informal sector x Female (1/0)	0.047**	0.029	0.092***	0.071***
	(0.013)	(0.015)	(0.019)	(0.017)
Female (1/0)	-0.225***	-0.225***	-0.177***	-0.177***
	(0.011)	(0.011)	(0.012)	(0.012)
Informal x Female x (Year = 2012)		0.015		0.024
		(0.014)		(0.015)
Informal x Female x (Year = 2013)		0.030*		0.034*
		(0.014)		(0.015)
Informal x Female x (Year = 2014)		0.010		0.012
		(0.015)		(0.012)
Informal x Female x (Year = 2015)		0.037*		0.035*
		(0.016)		(0.015)
Mean of dependent variable	6.936	6.936	6.988	6.988
Year and State FE	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
Observations	3.043e+08	3.043e+08	3.037e+08	3.037e+08

Note: Data are taken from PNAD 2011–15. The controls are sex, race, sector of employment, an interaction of sex and sector of employment, age, age squared, and tenure in current employment. Standard errors are clustered at the state level. Excluded are individuals who are self-employed, farm laborers, and domestic workers. Individuals are classified as working in the formal sector if they report that they have worked in the past week and have a formal contract with their employer. FE = fixed effect.

Table A3: Estimates with outcomes as rates or counts

	Homicides				Hospitalizations				Reports			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Leave-out F-M wage ratio	-1.1	-1.5	-3*	-1.7	-5.8	-8.1	-4.4	-3.2***	43	6.4	-4.5*	-1.6
	(4.9)	(4.9)	(1.7)	(1.6)	(8.4)	(10)	(4.6)	(1.2)	(52)	(51)	(2.7)	(1.4)
Model	Linear	Linear	Poisson	Poisson	Linear	Linear	Poisson	Poisson	Linear	Linear	Poisson	Poisson
Controls	N	Y	N	Y	N	Y	N	Y	N	Y	N	Y
Municipality FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Number of municipalities	841	841	821	821	841	841	744	744	841	841	751	751

Note: For the linear model, outcomes are in rates (per 100,000) women and for the Poisson model outcomes are in counts. The control variables are the natural log of the population of women ages 15–49, the number of households receiving cash transfers, divided by the population of women ages 15–49, the municipal per capita value of cash transfers, the ratio of females to males in secondary education completion, the share of the population ages 15–49 that is female, the per capita number of health clinics, and the per capita number of hospital beds. Regressions are weighted by the average number of women ages 15–49 in 2011–16. FE = fixed effect. F-M = female to male. Significance levels are reported as * $p < .10$ ** $p < .05$ *** $p < .01$.

Table A4: Main estimates, no population weights

	Homicides		Hospitalizations		Reports	
	(1)	(2)	(3)	(4)	(5)	(6)
Leave-out F-M wage ratio	-1.3	-1.4	.0081	.016	1.1	1.2
	(1.2)	(1.2)	(1.3)	(1.3)	(.86)	(.9)
Controls	N	Y	N	Y	N	Y
Municipality FE	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y
Number of municipalities	841	841	841	841	841	841

Note: The control variables are the natural log of the population of women ages 15–49, the number of households receiving cash transfers, divided by the population of women ages 15–49, the municipal per capita value of cash transfers, the ratio of females to males in secondary education completion, the share of the population ages 15–49 that is female, the per capita number of health clinics, and the per capita number of hospital beds. FE = fixed effect. F-M = female to male. Significance levels are reported as * $p < .10$ ** $p < .05$ *** $p < .01$.

Table A5: Results using employment as an instrument for the raw female-male wage ratio

	Homicides		Hospitalizations		Reports	
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: OLS						
F emp	-2.8**	-1.9	-2.2	-1.2	2.5*	2.7**
	(1.4)	(1.4)	(2.2)	(2.2)	(1.3)	(1.3)
M emp	2.2	1.4	4.1	3.3	-.12	-.5
	(1.3)	(1.3)	(3.2)	(3.1)	(1.3)	(1.3)
Panel B: First stage: predicting raw ratio w/ female and male employment						
F emp	.43***	.43***	.43***	.43***	.43***	.43***
	(.093)	(.09)	(.093)	(.09)	(.093)	(.09)
M emp	-.42***	-.42***	-.42***	-.42***	-.42***	-.42***
	(.1)	(.096)	(.1)	(.096)	(.1)	(.096)
Panel C: IV: Raw wage ratio instrumented by female and male employment						
Raw wage ratio (instrumented)	-6.2*	-4.1	-6	-3.6	4.6	5.6*
	(3.3)	(3.2)	(5.7)	(5.6)	(3.3)	(3.3)
First stage F-stat	10.70	11.20	10.70	11.20	10.70	11.20
Mean of dependent variable (natural log)	0.87	0.87	0.79	0.79	3.17	3.17
Controls	N	All	N	All	N	All
Municipality FE	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y
Number of municipalities	841	841	841	841	841	841

Note: The outcome variables are in natural logs; 0.01 is added to each value. The control variables are the natural log of the population of women ages 15–49, the number of households receiving cash transfers, divided by the population of women ages 15–49, the municipal per capita value of cash transfers, the ratio of females to males in secondary education completion, the share of the population ages 15– 49 that is female, the per capita number of health clinics, and the per capita number of hospital beds. Regressions are weighted by the average number of women ages 15–49 in 2011–16. FE = fixed effect. F-M = female to male. Significance levels are reported as * $p < .10$ ** $p < .05$ *** $p < .01$.

Table A6: Main results using the two most populous quintiles of municipalities

	Homicides		Hospitalizations		Reports	
	(1)	(2)	(3)	(4)	(5)	(6)
Leave-out F-M wage ratio	-1	-1.1*	-.2	-.081	-.65	-.48
	(.62)	(.63)	(.67)	(.67)	(.78)	(.77)
Controls	N	Y	N	Y	N	Y
Municipality FE	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y
Number of municipalities	1682	1682	1682	1682	1682	1682

Note: The outcome variables are in natural logs; 0.01 is added to each value. The control variables are the natural log of the population of women ages 15–49, the number of households receiving cash transfers, divided by the population of women ages 15–49, the municipal per capita value of cash transfers, the ratio of females to males in secondary education completion, the share of the population ages 15– 49 that is female, the per capita number of health clinics, and the per capita number of hospital beds. Regressions are weighted by the average number of women ages 15–49 in 2011–16. FE = fixed effect. F-M = female to male. Significance levels are reported as * p < .10 ** p < .05 *** p < .01.

Table A7: Main results with an alternative set of control variables

	Homicides			Hospitalizations			Reports		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Leave-out F-M wage ratio	-1.3	-1.3	-1.2	-.68	-.43	-.6	-.69	-.53	-.66
	(.89)	(.91)	(.9)	(1)	(1)	(1)	(1.2)	(1.2)	(1.1)
Ln F population [15-49]		.7	1.3		4.1**	3.6**		1.1	-.26
		(1.3)	(1.2)		(1.8)	(1.6)		(1.4)	(1.2)
Share of women w/Bolsa Familia		-.38			-1			-.4	
		(1.3)			(1.5)			(1.4)	
Per capita cash transfer		.016			-.0031			.011	
		(.01)			(.012)			(.011)	
F-M ratio with high school degree		-1.1*			-1.4			-.56	
		(.63)			(1.3)			(.84)	
Female share of population 15-49 yrs		-4.8			-34			-77***	
		(23)			(36)			(23)	
Clinics per 1,000 people		.022	.025		.024	.024		.039	.039
		(.021)	(.021)		(.029)	(.029)		(.024)	(.025)
Hospital beds per 1,000 people		-.012	-.012		-.0083	-.0068		-.0041	-.0045
		(.009)	(.0089)		(.014)	(.014)		(.01)	(.01)
Controls	N	Y	Y	N	Y	Y	N	Y	Y
Municipality FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
Number of municipalities	841	841	841	841	841	841	841	841	841

Note: Regressions are weighted by the average number of women ages 15–49 in 2011–16. FE = fixed effect. F-M = female to male. Significance levels are reported as * p < .10 ** p < .05 *** p < .01.

Table A8: Including State Trends

	Homicides		Hospitalizations		Reports	
	(1)	(2)	(3)	(4)	(5)	(6)
Leave-out F-M wage ratio	-1.3	-2.7**	-.43	.85	-.53	.27
	(.91)	(1.1)	(1)	(1.3)	(1.2)	(.81)
Controls	Y	Y	Y	Y	Y	Y
Municipality FE	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y
State time trend	N	Y	N	Y	N	Y
Number of municipalities	841	839	841	839	841	839

Note: The outcome variables are in natural logs; 0.01 is added to each value. The control variables are a one-year lag of the outcome variable, the natural log of the population of women ages 15–49, the number of households receiving cash transfers, divided by the population of women ages 15–49, the municipal per capita value of cash transfers, the ratio of females to males in secondary education completion, the share of the population ages 15–49 that is female, the per capita number of health clinics, and the per capita number of hospital beds. State trends are the state indicator variables interacted with a continuous year variable. Regressions are weighted by the average number of women ages 15–49 in 2011–16. FE = fixed e ff ect. F-M = female to male. Significance levels are reported as * p < .10 ** p < .05 *** p < .01.

Table A9: Placebo regressions using traffic accidents as outcomes

	Deaths		Hospitalizations	
	(1)	(2)	(3)	(4)
Leave-out F-M wage ratio	-.6	-.59	.055	-.33
	(1)	(.99)	(1.2)	(1.2)
Controls	N	Y	N	Y
Municipality FE	Y	Y	Y	Y
Year FE	Y	Y	Y	Y
Number of municipalities	841	841	841	841

Note: The outcome variables are in natural logs; 0.01 is added to each value. The control variables are the natural log of the population of women ages 15–49, the number of households receiving cash transfers, divided by the population of women ages 15–49, the municipal per capita value of cash transfers, the ratio of females to males in secondary education completion, the share of the population ages 15– 49 that is female, the per capita number of health clinics, and the per capita number of hospital beds. Regressions are weighted by the average number of women ages 15–49 in 2011–16. FE = fixed effect. F-M = female to male. Significance levels are reported as * p < .10 ** p < .05 *** p < .01.