Erratum: "Does import competition affect the gender gap? Evidence from matched employer-employee data"

The last sentence of the data description of the Veneto Worker History Panel dataset in the published version is incorrect. The VWH dataset actually covers all workers' employment spells in or out of Veneto, if that worker has worked at least a day in Veneto; this is the data used in this paper. The corrected description of the data should therefore be as follows:

The data covers all job spells of workers that ever worked at least a day in Veneto, including job spells outside Veneto; any workers that leave the sample may therefore be leaving because they exit the labour force or Italy.

The second footnote should also be slightly altered as a result to read as follows:

While this does not invalidate my findings highlighting the importance of worker and firm selection, I am nonetheless careful to state that workers are leaving the sample by exiting employment within Italy, and not exiting employment entirely. Note also that firms outside of Veneto are also included in the analysis, but only include employees that worked in Veneto at some point; this should also not invalidate the fixed effect and firm exit results as they pertain to employment for Veneto-related workers.

Finally, there were some coding and data compiling errors discovered, which meant that, for some of the first stage F-statistics and the results, there were slightly revised numbers. While the effect of import competition on the gender wage gap with only firm fixed effects is not statistically significant in the revised table 1, the results in the paper are largely not qualitatively affected otherwise, although significance and magnitudes vary slightly at times. One key exception is that the median worker effect results in Table 2 are now much weaker, mostly not maintaining their original significance. Table 2 does still show that women are disproportionately likely to leave the sample and the industry when impacted by import shocks, providing some additional support that selection and exit amongst women is stronger than for men. Table 3, which is largely unchanged qualitatively, also still shows that firms with more women are more likely to contract and exit when faced with an import shock. The revised Table 1 results still provide some support for the main

Table 1: Comparison of cross-sectional, person FE, and match FE results

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Method	OLS	OLS	OLS	OLS	IV	IV	IV	IV
Dep. Var.	$ln(w_{ifst})$	$ln(w_{ifst})$	$ln(w_{ifst})$	$ln(w_{ifst})$	$ln(w_{ifst})$	$ln(w_{ifst})$	$ln(w_{ifst})$	$ln(w_{ifst})$
IMP_{st}	-0.149	0.140	-0.162*	0.409**	-2.088	0.273*	-1.725	0.306
	(0.111)	(0.091)	(0.092)	(0.157)	(1.680)	(0.138)	(1.329)	(0.483)
IMP_{st} x $female_i$	0.137	-0.581***	0.089	-0.772***	-0.281	-2.046***	-0.254	-3.896***
	(0.103)	(0.123)	(0.092)	(0.209)	(0.208)	(0.574)	(0.187)	(1.242)
industry FE	yes	yes	no	no	yes	yes	no	no
worker FE	no	yes	no	no	no	yes	no	no
firm FE	no	no	yes	no	no	no	yes	no
match FE	no	no	no	yes	no	no	no	yes
F-stat. (IMP_{st})					3.96	31.51	4.51	6.46
F-stat. (IMP_{st} x $female_i$)					16.90	7.90	18.89	2.33
Observations	8,542,112	8,542,112	8,542,112	8,542,112	8,542,112	8,321,250	8,537,453	7,883,545
R-squared	0.459	0.681	0.551	0.739	0.362	0.146	0.300	0.080

*** denotes 1% significance, ** denotes 5% significance, * denotes 10% significance. Robust standard errors, clustered by industry, are reported in parentheses. All columns include age_{ifst}, age_{ifst}^2, tenure$_{ifst}$, and $tenure_{ifst}^2$ and occupation and year fixed effects as control variables. Columns 1, 3,5, and 7 also include a gender dummy variable.

mechanism of the paper, that sorting and selection of women can be concealed by estimating only cross-sectional regressions. The fully revised tables are presented below:

Table 2: Industry switchers' and sample leavers' characteristics

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Method	OLS	OLS	IV	IV	OLS	OLS	IV	IV
Sample	sample leavers	sample leavers	sample leavers	sample leavers	ind. switchers	ind. switchers	ind. switchers	ind. switchers
Dep. Var.	ln(leavers)	median worker FE	ln(leavers)	median worker FE	ln(leavers)	median worker FE	ln(leavers)	median worker FE
IMP_{st}	-3.782***	0.058	-6.472*	0.152	-3.313***	-0.067*	-5.296*	0.103
	(1.294)	(0.049)	(3.411)	(0.130)	(1.187)	(0.039)	(3.054)	(0.121)
IMP_{st} x $female_g$	7.537***	-0.036	12.87***	-0.005	7.597***	-0.088**	12.70***	-0.039
	(2.409)	(0.0454)	(4.077)	(0.0584)	(2.312)	(0.0348)	(3.821)	(0.055)
Industry FE	yes	yes	yes	yes	yes	yes	yes	yes
Gender FE	yes	yes	yes	yes	yes	yes	yes	yes
Year FE	yes	yes	yes	yes	yes	yes	yes	yes
Observations	2,651	2,627	2,651	2,627	2,679	2,660	2,679	2,660
R-squared	0.114	0.156	0.485	0.183	0.133	0.199	0.508	0.111

*** denotes 1% significance, ** denotes 5% significance, * denotes 10% significance. Robust standard errors, clustered by industry, are reported in parentheses. Observations are at the industry-gender-year level. The dependent variable in column 1 is the logged number of workers who leave the industry for employment in another industry, for the industry-gender-year. The dependent variable in column 2 is the median worker fixed effect of workers who switch industries from the industry-gender-year. The dependent variable in column 3 is the logged number of workers that leave the sample from the industry-gender-year cell. The dependent variable in column 4 is the median worker fixed effect from workers who leave the sample, for the given industry-gender-year cell.

Table 3: Firm size and exit

	(1)	(2)	(3)	(4)
Method	OLS	OLS	IV	IV
Dep. Var.	$ln(emp_{fst})$	$exit_{fst}$	$ln(emp_{fst})$	$exit_{fst}$
IMP_{st}	0.449	-0.292***	0.665	-0.0032
	(0.295)	(0.138)	(0.595)	(0.381)
$fem.share_{fst}$	0.065	0.0383***	0.213*	-0.0245
	(0.0522)	(0.0121)	(0.111)	(0.0381)
IMP_{st} x $fem.share_{fst}$	-1.122***	0.580***	-5.187***	2.378***
	(0.392)	(0.178)	(1.596)	(0.733)
white collar share$_{fst}$	-0.494***	-0.0792***	-0.526***	-0.0633***
	(0.0453)	(0.0133)	(0.0547)	(0.0107)
$ln(emp_{fst})$		-0.168***		-0.166***
		(0.00964)		(0.0086)
Firm FE	yes	yes	yes	yes
Year FE	yes	yes	yes	yes
Observations	329,856	308,579	323,151	302,187
R-squared	0.073	0.0254	0.0425	0.0925

*** denotes 1% significance, ** denotes 5% significance, * denotes 10% significance. Robust standard errors, clustered by industry, are reported in parentheses.