## Appendix

## Robustness checks

## Replace the PNTR measure with data from Pierce and Schott (2016)

 Since PNTR is a constructed measure of regional trade exposure rather than a directly recorded measure in the data, I thereby try to reconstruct it using an alternative data source. Pierce and Schott (2016) provided similar tariff reduction data as that of Erten and Leight (2021), but at more finely-defined industrial scopes, one that was also used by Facchini et al. (2019) when studying the migration behavior of Chinese migrants. I follow Pierce and Schott (2016) and Facchini et al. (2019) to reconstruct the PNTR measure and re-regress equation 3 in the main text. Following Dai et al. (2021), I drop workers in the agriculture sector rather than treating them as a zero number to save efficiency. Table A1 presents the estimation results. As the data miss some observations of the experience variable in certain years, the resulting sample size of the regression is smaller when the experience variable is included. The coefficients on *male\_pntr\_post* vary from 0.048 to 0.051 and are all positive and statistically significant at the 1% level, which is somewhat larger than that of the baseline estimates using the Erten and Leight (2021) data.

[TABLE A1 HERE]

## Controlling additional workers’ characteristics

[TABLE A2 HERE]

Table A2 further presents the estimation results of equation 3 in the main text while controlling additional personal demographic and socioeconomic characteristics such as education, industry, occupation, and ownership structure of the employer. The main results remain unchanged, estimations of $β\_{1}$ are all positive and statistically and economically significant at conventional levels. Moving across the columns from left to right in Table A2 shows that the estimate for $β\_{1}$ decreases in absolute value as additional control variables are included.

## Rerun the specification at the aggregated level

 This subsection additionally tests the robustness of the estimates by aggregating the individual wage data to the province level to construct a province-level gender wage gap measure and using the obtained gender wage gap measure to estimate a difference-in-difference (DID) regression. Specifically, I estimate two regression equations expressed in equation A1 and equation A2. Equation A1 is a continuous DID regression specification, with the explanatory variable being the province-level PNTR measure and the dependent variable being the province-level gender wage gap. It includes a set of province control variables and a battery of fixed effects. $gap\_{pt}$ represents the gender wage gap of province $p$ in year $t$, while $pntr\_{p,1999}$ is the province-level measure of trade exposure.

$ ln\left(gap\_{pt}\right)=β\_{0}+β\_{1}\*pntr\_{p,1999}\*post\_{t}+γ\_{p}+γ\_{t}+ε\_{pt}$ (A1)

 Equation A2 is a discrete form of the DID regression specification, with the explanatory variable being a binary variable of whether a province is in high trade exposure or low trade exposure. Provinces are classified as high or low according to their relative size of the PNTR measure to the median PNTR. Specification of this sort is akin to that of Han et al. (2012) who split provinces into high-exposed and low-exposed regions when studying the impact of globalization on wage inequality in urban China.

 $ln\left(gap\_{pt}\right)=β\_{0}+β\_{1}high\_{p}\*post\_{t}+γ\_{p}+γ\_{t}+ε\_{pt}$ (A2)

$high\_{p}$ is a dummy variable that indicates whether province $p$ is a high-exposed region or not, while $post\_{t}$ indicates whether year $t$ is in post-2001 period or not. $γ\_{p}$ is the province fixed effect, $γ\_{t}$ is the year fixed effect, and $ε\_{pt}$ is a random error term. The coefficient of interest is $β\_{1}$.

[TABLE A3 HERE]

Table A3 reports the regression results of equation A1. Positive coefficients on the *pntr\_post* variable in all columns indicate that high-exposed regions experience larger increases in the gender wage gap. The coefficients in columns 1 and 2 are both significant at the 5% level. Key findings remain unchanged when measuring the gender wage gap in its natural logarithmic form, as seen in columns 3 and 4.

[TABLE A4 HERE]

Table A4 reports the estimation results of equation A2, the results are similar to that in Table A3, even if the coefficients are larger. Table A3 and Table A4 both confirm the finding that trade liberalisation causes a rise in the gender wage gap.

## Test the pre-policy parallel trend assumption

**FigureA1** Event study graph of multiple DDD regression

Figure A1 provides graphical event-study evidence on the effectiveness of the pre-policy parallel trend hypothesis between high and low trade-exposed regions and provides year-specific coefficients on the *male\_pntr\_post* variable for 18 consecutive years spanning the 1992-2009 period, drawing data from Erten and Leight (2021). There’s a clear rise in the gender wage gap after the WTO accession event. The coefficient jumps from 2001 to 2002 and is significantly different from zero after 2002, while it is indistinguishable from zero before 2002. The pattern shown in Figure A1 indicates that the parallel trend hypothesis among high and low trade-exposed regions before the trade shock indeed holds.

## References

1. **Dai M, Huang W and Zhang Y** (2021) How Do Households Adjust to Tariff Liberalisation? Evidence from China’s WTO Accession. *Journal of Development Economics* **150**(3), 102628. <https://doi.org/10.1016/j.jdeveco.2021.102628>.
2. **Erten B and Leight J** (2021) Exporting Out of Agriculture: The Impact of WTO Accessionon Structural Transformation in China. *Review of Economics and Statistics* **103**(2), 364-380. <https://doi.org/10.1162/rest_a_00852>.
3. **Facchini G, Liu M, Mayda A and Zhou M** (2019) China’s “Great Migration”: The Impact of the Reduction in Trade Policy Uncertainty. *Journal of International Economics* **120**(5), 126-144. <https://doi.org/10.1016/j.jinteco.2019.04.002>.
4. **Han J, Liu R and Zhang J** (2012) Globalisation and Wage Inequality: Evidence from Urban China. *Journal of International Economics*, **87**(2), 288-297.

<https://doi.org/10.1016/j.jinteco.2011.12.006>.

1. **Pierce J and Schott P** (2016) The Surprisingly Swift Decline of US Manufacturing Employment. *American Economic Review* **106**(7), 1632-1662.

http://dx.doi.org/10.1257/aer.20131578.

## Additional Tables and Figures

 **Table A1** Regression results using data from Pierce and Schott (2016)

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | Column 1 | Column 2 | Column 3 | Column 4 | Column 5 | Column 6 |
|  | lnwage | lnwage | lnwage | lnwage | lnwage | lnwage |
| male\_pntr\_post | 0.051\*\*\* | 0.049\*\*\* | 0.050\*\*\* | 0.048\*\*\* | 0.049\*\*\* | 0.050\*\*\* |
|  | (0.01) | (0.01) | (0.02) | (0.01) | (0.01) | (0.01) |
| Age |  | 0.073\*\*\* |  | 0.026\*\*\* | 0.082\*\*\* | 0.030\*\*\* |
|  |  | (0.00) |  | (0.00) | (0.00) | (0.00) |
| Age sq. |  | -0.001\*\*\* |  | -0.000\*\*\* | -0.001\*\*\* | -0.000\*\*\* |
|  |  | (0.00) |  | (0.00) | (0.00) | (0.00) |
| Experience |  |  | 0.035\*\*\* | 0.028\*\*\* |  | 0.032\*\*\* |
|  |  |  | (0.00) | (0.00) |  | (0.00) |
| Exp. Sq. |  |  | -0.001\*\*\* | -0.001\*\*\* |  | -0.001\*\*\* |
|  |  |  | (0.00) | (0.00) |  | (0.00) |
| Education |  |  |  |  | Yes | Yes |
| N | 347119 | 347119 | 337386 | 337386 | 347119 | 337386 |
| adj. *R*2 | 0.190 | 0.204 | 0.210 | 0.212 | 0.320 | 0.330 |

Notes: Data are drawn from the 1992-2009 UHS microdata and are matched with the Pierce and Schott (2016) tariff reduction data; All specifications include province-by-year, male-by-province, and male-by-year fixed effects; Standard errors in parentheses are clustered at the province-by-year level; \* p < 0.1, \*\* p < 0.05, \*\*\* p < 0.01.

**Table A2** Controlling workers’ industries and occupations

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | Column 1 | Column 2 | Column 3 | Column 4 | Column 5 | Column 6 | Column 7 |
|  | lnwage | lnwage | lnwage | lnwage | lnwage | lnwage | lnwage |
| male\_pntr\_post | 0.028\*\*\* | 0.024\*\* | 0.022\*\* | 0.021\*\* | 0.014\* | 0.013\* | 0.015\*\* |
|  | (0.010) | (0.010) | (0.010) | (0.008) | (0.008) | (0.008) | (0.007) |
| Age |  | Yes | Yes | Yes | Yes | Yes | Yes |
| Experience |  |  | Yes | Yes | Yes | Yes | Yes |
| Education |  |  |  | Yes | Yes | Yes | Yes |
| Industry  |  |  |  |  | Yes | Yes | Yes |
| Occupation |  |  |  |  |  | Yes | Yes |
| Ownership |  |  |  |  |  |  | Yes |
| *N* | 395964 | 395964 | 384450 | 384450 | 384125 | 384125 | 384125 |
| adj. *R*2 | 0.211 | 0.229 | 0.235 | 0.362 | 0.393 | 0.404 | 0.416 |

Notes: The PNTR measure is calculated using industrial NTR data from Erten and Leight (2021); All columns include the year, province, male-by-year, male-by-province, and province-by-year fixed effects; Standard errors in parentheses are clustered at the province-by-year level; \* p < 0.1, \*\* p < 0.05, \*\*\* p < 0.01.

**Table A3** Regressions at the aggregate level

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Column 1 | Column 2 | Column 3 | Column 4 |
|  | gap | lngap | gap | lngap |
| pntr\_post | 0.024\*\* | 0.018\*\* | 0.024 | 0.018 |
|  | (0.01) | (0.01) | (0.03) | (0.02) |
| *N* | 273 | 273 | 273 | 273 |
| adj. *R*2 | 0.619 | 0.606 | 0.619 | 0.606 |

Notes: Standard errors in parentheses, \* p < 0.1, \*\* p < 0.05, \*\*\* p < 0.01; Robust standard errors in columns (1) and (2), standard errors are clustered at the province level in columns (3) and (4); The province and year fixed effects are included in all specifications.

**Table A4** Regression results using binary treatment variables with aggregated data

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Column 1 | Column 2 | Column 3 | Column 4 |
|  | gap | lngap | gap | lngap |
| high\_post | 0.048\*\*\* | 0.040\*\*\* | 0.048 | 0.040 |
|  | (0.02) | (0.01) | (0.04) | (0.03) |
| *N* | 291 | 291 | 291 | 291 |
| adj. *R*2 | 0.637 | 0.626 | 0.637 | 0.626 |

Notes: Standard errors in parentheses; Robust standard errors in columns (1) and (2), standard errors are clustered at the province level in columns (3) and (4); Province and year fixed effects are included in all specifications. Specifically, the high-exposed regions are provinces where PNTR measures are above the median PNTR while the remaining provinces are low-exposed regions. The high-exposed regions include Beijing, Guangdong, Zhejiang, Jiangsu, Liaoning, Hubei, Shandong, Heilongjiang, and Shanghai, while Gansu, Shaanxi, Guizhou, Yunnan, Sichuan, Henan, Hunan, Shanxi, and Jiangxi are low-exposed regions; \* p < 0.1, \*\* p < 0.05, \*\*\* p < 0.01.

 **FigureA1** Event study graph of multiple DDD regression