Wage Variations and Commuting Distance

Clément Nedoncelle, El-Mehdi Aboulkacem

Online Appendix – Not for Publication

Wage Variations and Commuting Distance

El-Mehdi Aboulkacem [1]

Clément Nedoncelle *[2]

[1] UGE – AME - SPLOTT, 59666, Villeneuve d'Ascq Cedex, France
[2] Université Paris-Saclay, INRAE, AgroParisTech, Paris-Saclay Applied Economics, 91120, Palaiseau, France.

Appendix A: Additional Results

Table A1 – Domestic Sales – Short-run

Dependent Var.: Log Commuting Distance	(1)	(2)	(3)	(4)	(5)
Log Yearly Net Wage	0.156***	0.189***	0.167***	0.200***	0.194***
	(0.014)	(0.017)	(0.015)	(0.019)	(0.023)
Log Domestic Sales	-0.011***	0.011***	0.010***	0.010***	0.008***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Worker FE	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓
Firm-Year Controls		✓		✓	✓
Worker Controls			✓	✓	✓
Occupation FE					✓
Observations	1802792	1802792	1802792	1802792	1802792

Notes: Second-stage IV estimations, in which World Import Demand (see text) is used as excluded instrument for yearly net wages. Robust standard errors in parentheses with ***, **, and *respectively denoting significance at the 1%, 5% and 10% levels. Standard errors are clustered at the sector-year level. Firm-Year controls include log Assets, Log Apparent Labor Productivity, defined as total value added per worker, and log Employment. Worker controls include a skilled/unskilled dummy, the age of the worker and gender. Domestic Sales from the balance-sheet data.

Table A2 – Domestic Sales – Long-run

Dependent Var.: Log Commuting Distance					
	(1)	(2)	(3)	(4)	(5)
Log Yearly Net Wage	0.218***	0.263***	0.242***	0.285***	0.304***
	(0.023)	(0.039)	(0.027)	(0.043)	(0.065)
Log Domestic Sales	-0.010***	-0.013***	-0.011***	0.014***	-0.014***
	(0.001)	(0.002)	(0.001)	(0.002)	(0.001)
Worker FE	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓
Firm-Year Controls		✓		✓	✓
Worker Controls			✓	✓	✓
Occupation FE					✓
Observations	280822	280822	280822	280822	280822

Notes: Second-stage IV estimations, in which World Import Demand (see text) is used as excluded instrument for yearly net wages. Robust standard errors in parentheses with ***, **, and *respectively denoting significance at the 1%, 5% and 10% levels. Standard errors are clustered at the sector-year level. Firm-Year controls include log Assets, Log Apparent Labor Productivity, defined as total value added per worker, and log Employment. Worker controls include a skilled/unskilled dummy, the age of the worker and gender. Domestic Sales from the balance-sheet data.

Table A3 – Previous Export Sales – Short-run

Dependent Var.: Log Commuting Distance	(1)	(2)	(3)	(4)	(5)
Log Yearly Net Wage	0.193***	0.214***	0.202***	0.224***	0.225***
	(0.022)	(0.025)	(0.023)	(0.027)	(0.023)
Log Exports *t-1*	0.001	0.001	0.002**	0.002***	0.002***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Log Nb HS6 products *t-1*	-0.014***	-0.012***	-0.013***	-0.011***	-0.011***
	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)
Worker FE	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓
Firm-Year Controls		✓		✓	✓
Worker Controls			✓	✓	✓
Occupation FE					✓
Observations	1310012	1310012	1310012	1310012	1310012

Notes: Second-stage IV estimations, in which World Import Demand (see text) is used as excluded instrument for yearly net wages. Robust standard errors in parentheses with ***, **, and *respectively denoting significance at the 1%, 5% and 10% levels. Standard errors are clustered at the sector-year level. Firm-Year controls include log Assets, Log Apparent Labor Productivity, defined as total value added per worker, and log Employment. Worker controls include a skilled/unskilled dummy, the age of the worker and gender. Exports-related variables from the customs data.

Table A4 – Previous Export Sales – Long-run

Dependent Var.: Log Commuting Distance	(1)	(2)	(3)	(4)	(5)
Log Yearly Net Wage	0.253***	0.305***	0.264***	0.322***	0.349***
	(0.022)	(0.043)	(0.038)	(0.048)	(0.069)
Log Exports *t-1*	0.006	0.003	-0.002	0.001	0.003*
	(0.001)	(0.001)	(0.001)	(0.004)	(0.002)
Log Nb HS6 products *t-1*	-0.007***	-0.012**	-0.013***	-0.017***	-0.019***
	(0.007)	(0.005)	(0.005)	(0.004)	(0.003)
Worker FE	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓
Firm-Year Controls		✓		✓	✓
Worker Controls			✓	✓	✓
Occupation FE					✓
Observations	280822	280822	280822	280822	280822

Notes: Second-stage IV estimations, in which World Import Demand (see text) is used as excluded instrument for yearly net wages. Robust standard errors in parentheses with ***, **, and *respectively denoting significance at the 1%, 5% and 10% levels. Standard errors are clustered at the sector-year level. Firm-Year controls include log Assets, Log Apparent Labor Productivity, defined as total value added per worker, and log Employment. Worker controls include a skilled/unskilled dummy, the age of the worker and gender. Exports-related variables from the customs data.

Appendix B: Plausibly exogenous IV

The validity of our IV results relies on the validity of the excluded instrument we used. We now check the validity of the world import demand as an instrument and assess the sensitivity of the second-stage results. We proceed in 4 steps.

Observables and Excluded Instrument

Our IV strategy uses sources of variations in wages that should be uncorrelated to the firms and to workers' observables to allow for causal inference. Our excluded instrument should in particular be uncorrelated to firms' and workers' observables.

In Table B1 – Panel A, we show that, altogether, workers' observables cannot explain the excluded instrument. In these estimations, we regress the observed instrument on the full set of workers' observables, and we obtain low f-statistics on these regressions: all f-stats are under the standard confidence levels, in particular as we have many observations. Then, Table B2 shows that there is no strong correlation between our baseline instrument and the set of workers' observables. In particular, we hardly estimate a significant and robust correlation independently of the fixed effects we use. These two results suggest that our instrument is plausibly exogenous to workers: the instrument is quasi-randomly distributed across workers in our sample.

Table B1 – Panel B and Table B3 show there is no strong correlation between the excluded instrument and the firm controls either. Indeed, regressions of firm observables on the world import demand provide hardly significant coefficients. As for workers, the instrument is quasi-randomly distributed across firms in our sample. If one worries one might worry about bias from unobserved firm attributes, Table B4 hereafter precisely accounts for this potential bias. Maybe firms are not completely randomly facing random foreign shocks: there may be some unobserved firm attributes that lead some firms to be more exposed to shocks than others. We follow Borusyak and Hull (2020) to exclude this bias. For details, please refer to the discussion hereafter.

Table B1 – Instrument and Workers and Firms Observables

Panel A: Workers' Observables

Dependent Var.: Log WID (Current Weights)					
	(1)	(2)	(3)	(4)	(5)
Age	0.150			0.078	-0.040
	(0.231)			(0.225)	(0.207)
Gender		0.092		0.121	0.244***
		(0.067)		(0.069)	(0.066)
Skilled Dummy			0.281***	0.286***	
			(0.098)	(0.089)	
F-stat	0.422	1.891	8.177	3.989	10.702
Year FE	✓	✓	✓	✓	✓
Occupation FE					✓
Observations	1597288	1597288	1597288	1597288	1597288

Panel B: Firms' Observables

Dependent Var.: Log WID (Current Weights)						
	(1)	(2)	(3)	(4)	(5)	(6)
Assets	0.120**				0.544***	0.220**
	(0.058)				(0.132)	(0.101)
Labor Prod.		0.568***			0.250	0.164*
		(0.183)			(0.163)	(0.097)
Employment			0.058		-0.344***	-0.069
			(0.052)		(0.097)	(0.090)
Total Sales				0.040	-0.110*	-0.043
				(0.044)	(0.062)	(0.045)
F-stat	4.31	9.61	1.21	0.84	7.64	6.94
Worker FE						✓
Year FE	✓	✓	✓	✓	✓	✓
Occupation FE					✓	✓
Observations	1597288	1597288	1597288	1597288	1597288	1597288

Notes: Robust standard errors in parentheses with ***, **, and *respectively denoting significance at the 1%, 5% and 10% levels. Standard errors are clustered at the sector-year level.

Table B2 – Workers' Observables and Excluded Instrument – Short-run

Dependent Variable:	Log Age		Gender Dummy		Skilled Dummy	
	(1)	(2)	(3)	(4)	(5)	(6)
Log WID (Current Weights)	-0.002	0.000	0.002	0.001	0.003	0.002*
	(0.001)	(0.000)	(0.003)	(0.001)	(0.003)	(0.001)
Worker FE	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓
Firm-Year Controls	✓	✓	✓	✓	✓	✓
Worker Controls	✓	✓	✓	✓	✓	✓
Occupation FE		✓		✓		✓
Observations	1597288	1597288	1597288	1597288	1597288	1597288

Notes: Robust standard errors in parentheses with ***, **, and *respectively denoting significance at the 1%, 5% and 10% levels. Standard errors are clustered at the sector-year level. Firm-Year controls include log Assets, Log Apparent Labor Productivity, defined as total value added per worker, and log Employment. Worker controls include a skilled/unskilled dummy, the age of the worker and gender. The variable that is considered as the dependent variable is excluded from the controls.

Table B3 – Firm-level Observables and Excluded Instrument -Short-run

Dependent Variable:	Log Assets		Log Labor Prod.		Log Employment	
	(1)	(2)	(3)	(4)	(5)	(6)
Log WID (Current Weights)	-0.004***	0.006***	0.0065	0.005	0.007*	-0.003*
	(0.002)	(0.004)	(0.003)	(0.004)	(0.004)	(0.002)
Worker FE		✓		✓		✓
Year FE	✓	✓	✓	✓	✓	✓
Firm-Year Controls	✓	✓	✓	✓	✓	✓
Worker Controls	✓	✓	✓	✓	✓	✓
Occupation FE		✓		✓		✓
Observations	1597288	1597288	1597288	1597288	1597288	1597288

Notes: Robust standard errors in parentheses with ***, **, and *respectively denoting significance at the 1%, 5% and 10% levels. Standard errors are clustered at the sector-year level. Firm-Year controls include log Assets, Log Apparent Labor Productivity, defined as total value added per worker, and log Employment. Worker controls include a skilled/unskilled dummy, the age of the worker and gender. The variable that is considered as the dependent variable is excluded from the controls.

Non-random exposure of firms to foreign shocks.

By construction, a shift-share instrument combines foreign exogenous shocks with firm-specific exposure to these shocks. Intuitively, foreign shocks can hardly be correlated to firm and workers' observables. Yet, a remaining concern is that firms are non-randomly exposed to those random export shocks. For instance, we could imagine that some firms have an export structure which is oriented toward specific foreign markets which may be correlated to the employment of specific workers (for instance, younger and more prone to live further away from their firm). In other words, the non-random exposure to the random shock may lead to a bias in our estimations. Our results should thus be insulated from this non-random exposure of firms to foreign demand shocks.

Borusyak & Hull (2020) develop a general econometric framework for shift-share settings that allow for the possibility that shock exposure is non-random. They also provide a set of guidelines to implement this strategy in empirical applications. We follow these guidelines and use a "re-centering" process, i.e. we control for the random average exposure of firms, based on unobservables.

In formal terms, we have computed a set of 100 random counterfactual country-product-year level shocks, replicating the average distribution of the observed shocks (M_{jst}). Armed with these 100 random shifts, we combined them with the observed firm-level weights, creating a set of 100 counterfactual instruments (i.e. firm-level trade shifts), then averaged by firm-year. We thus computed an average exposure of a firm to foreign shocks, out of random shocks. Following the idea in Borusyak & Hull (2020), this *average instrument* should capture the differential (and potential non-random) exposure of the firm to random shocks.[1]

We include this average firm-year instrument in the second-stage estimation. Results are presented in Table B4 and Table B5. Overall, controlling for the average, random exposure of firms and workers to foreign shocks, we obtain close estimates to our baseline results. Then, we also do not estimate a significant impact of the correction we implement here. Our instrument thus does a good job at capturing the demand shifts abroad, without being contaminated by non-random exposure of firms.

[1] Intuitively, suppose for instance that firm A is more exposed to foreign shocks than firm B, both belonging to the same sector or industry, only because firm A is more centrally located. Then, the centrality of firm A is likely to be captured in the average instrument. Non-random centrality should thus contaminate each instrument and thus the average instrument. Controlling for this average instrument thus acts as controlling for non-random centrality in this example.

Table B4 - Accounting for the non-random exposure of firms to foreign shocks – Short-run

Dependent Var.: Log Commuting Distance	(1)	(2)	(3)	(4)	(5)
Log Yearly Net Wage	0.142***	0.192***	0.154***	0.212***	0.198***
	(0.015)	(0.016)	(0.017)	(0.021)	(0.023)
Log Expected Instrument	0.001	-0.002**	0.001	-0.002**	-0.002**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Worker FE	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓
Firm-Year Controls		✓		✓	✓
Worker Controls			✓	✓	✓
Occupation FE					✓
Observations	1792162	1792162	1792162	1792162	1792162

Notes: Second-stage IV estimations, in which World Import Demand (see text) is used as excluded instrument for yearly net wages. Robust standard errors in parentheses with ***, **, and *respectively denoting significance at the 1%, 5% and 10% levels. Standard errors are clustered at the sector-year level. Firm-Year controls include log Assets, Log Apparent Labor Productivity, defined as total value added per worker, and log Employment. Worker controls include a skilled/unskilled dummy, the age of the worker and gender. The expected Instrument variable is randomly created to capture the differential exposure of firms and workers to foreign demand shocks (see main text for details).

Table B5 - Accounting for the non-random exposure of firms to foreign shocks. - Long-run

Dependent Var.: Log Commuting Distance					
	(1)	(2)	(3)	(4)	(5)
Log Yearly Net Wage	0.324***	0.471***	0.377***	0.507***	0.614***
	(0.022)	(0.047)	(0.032)	(0.058)	(0.094)
Log Expected Instrument	-0.006***	-0.008***	-0.007***	-0.007***	-0.007***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Worker FE	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓
Firm-Year Controls		✓		✓	✓
Worker Controls			✓	✓	✓
Occupation FE					✓
Observations	276502	276502	276502	276502	276502

Notes: Second-stage IV estimations, in which World Import Demand (see text) is used as excluded instrument for yearly net wages. Robust standard errors in parentheses with ***, **, and *respectively denoting significance at the 1%, 5% and 10% levels. Standard errors are clustered at the sector-year level. Firm-Year controls include log Assets, Log Apparent Labor Productivity, defined as total value added per worker, and log Employment. Worker controls include a skilled/unskilled dummy, the age of the worker and gender.

Plausibly exogenous instrument and resulting bounds

Another way to see the identification issues is to consider that the exclusion restriction is likely to hold in our exercise, but that we cannot exclude some minor violations of this condition. We thus implement a methodology proposed by Conley et al. (2012) to take into account this potential violation of exclusion restrictions. Intuitively, it consists in assessing to which extent the parameter of interest is actually biased if the exclusion restriction is close but not exactly satisfied. We provide the graphical representation emphasized in Conley et al. (2012).

In the short-run, Figure B1 shows that the estimated α_1 lies between 0.05 and 0.2, slightly depending on the size of the direct effect of world import demand on commuting distance (which should be exactly 0 if exclusion restrictions were perfectly met). Therefore, the IV results appear stable, despite potential minor violations of the exclusion restriction. In the long run, Figure B2 suggests that the estimated elasticities range between 0.1 and 0.3, i.e. twice as large as in the short run.

Figure B1 – Resulting bounds for α_1 - Conley et al. (2012) – Short-run estimates

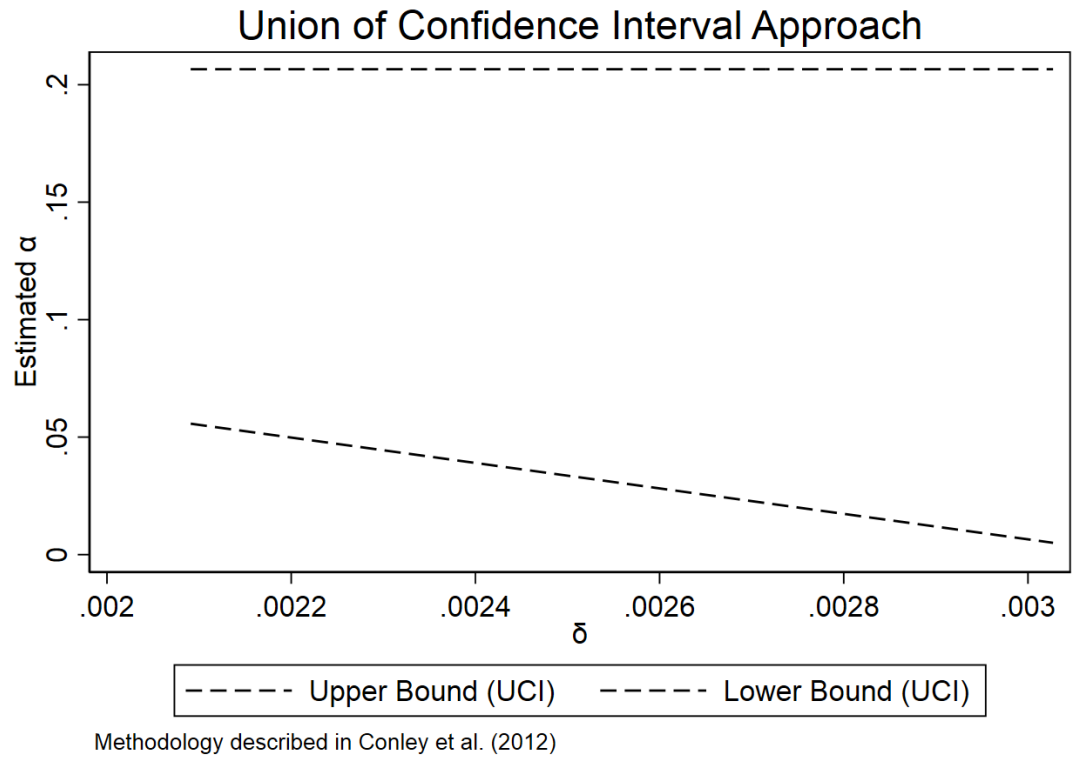
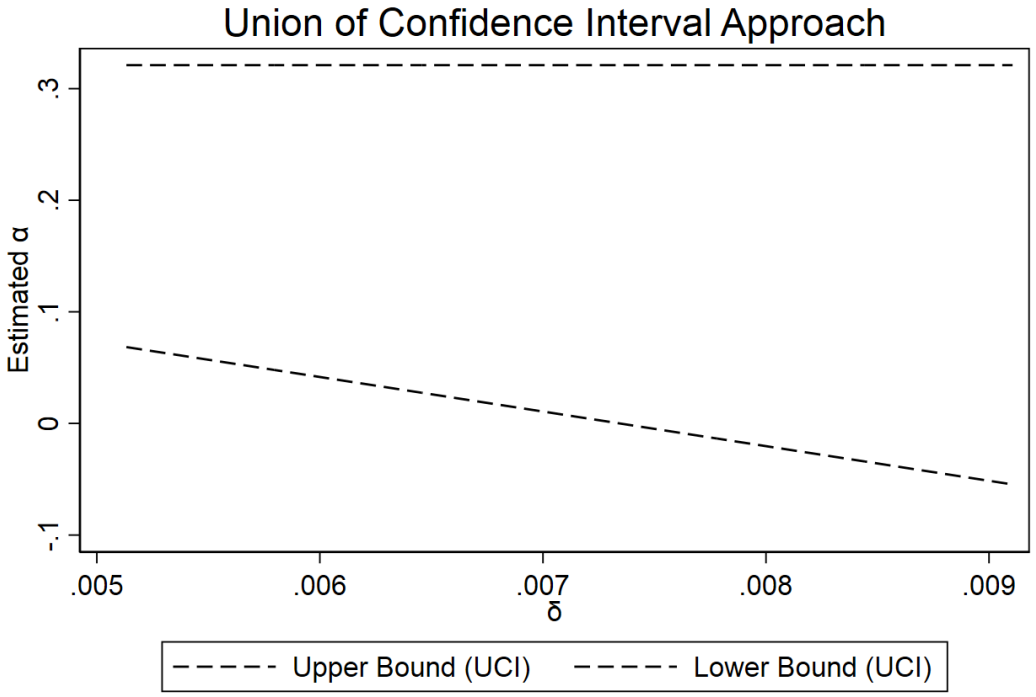


Figure B2 - Resulting Bounds for α_1 - Conley et al. (2012) – Long-run estimates



Methodology described in Conley et al. (2012)

Overidentifying assumptions

Finally, we depart from our baseline estimations by including a second excluded instrument in the analysis. We use the alternative measures of foreign product-level demand shocks as additional excluded instruments in our analysis. Many works in the literature (following Autor et al. (2013) for instance) 10-year changes in shocks as the identifying variation, instead of the contemporaneous levels. We follow this intuition and consider a 10-year change in demand shocks:

$$WID_{ft}^{10y-growth} = \sum_{js} \omega_{fjst} \times \Delta_{10} M_{jst}.$$

Our first stage estimation now becomes

$$\ln W_{ift} = \beta_1 \ln WID_{ft} + \beta_2 WID_{ft}^{10y-growth} + \beta_3 C_{ft} + \beta_4 C_{it} + FE + \varepsilon_{ift}^{first-stage}$$

while our second stage specification is unchanged.

Tables B6 and B7 present the estimation results and statistics to check the validity of our instruments. These replicate the structure of Table 2 in terms of dependent variable and specification, but now include two IVs. Including a second IV allows us to test for overidentifying restrictions. We check the validity of our instruments using two different tests. Robust to heteroskedasticity and clustering, Hansen J statistics for overidentifying restrictions are unable to reject our set of instruments. The F-stat form of the Kleibergen-Paap statistic as a test for weak instruments is also reported. All statistics are well above the critical values, confirming that our choice of instruments is statistically appropriate, on top of intuitive reasons explained above.

Second-stage results are close to our baseline estimates. We always find a positive and significant relationship between yearly wage and commuting distance, of the same magnitude (around 0.15 in the short run, and 0.3 in the long run) as in the baseline estimates. Note that the precision of these estimates is higher, mainly because we included a second IV in the first stage.

Table B6 – Overidentifying restrictions – Full Sample, Short-run estimates

Dependent Var.: Log Commuting Distance					
	(1)	(2)	(3)	(4)	(5)
Log Yearly Net Wage	0.142***	0.166***	0.153***	0.180***	0.146***
	(0.013)	(0.016)	(0.017)	(0.020)	(0.022)
Worker FE	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓
Firm-Year Controls		✓		✓	✓
Worker Controls			✓	✓	✓
Occupation FE					✓
Observations	1671808	1671808	1671808	1671808	1671808
Kleibergen-Paap rk Wald F statistic	678.965	521.997	465.522	354.029	281.461
Hansen J statistic	0.312	5.295	0.238	1.188	3.816
Hansen J statistic p-value	0.576	0.021	0.626	0.276	0.051

Notes: Second-stage IV estimations, in which World Import Demand with both contemporaneous and 10-year change in demand measures (see text) are used as excluded instruments for yearly net wages. Robust standard errors in parentheses with ***, **, and *respectively denoting significance at the 1%, 5% and 10% levels. Standard errors are clustered at the sector-year level. Firm-Year controls include log Assets, Log Apparent Labor Productivity, defined as total value added per worker, and log Employment. Worker controls include a skilled/unskilled dummy, the age of the worker and gender.

Table B7 – Overidentifying restrictions - Movers Sample, Long-run estimates

Dependent Var.: Log Commuting Distance					
	(1)	(2)	(3)	(4)	(5)
Log Yearly Net Wage	0.192***	0.259***	0.232***	0.311***	0.286***
	(0.024)	(0.036)	(0.029)	(0.041)	(0.060)
Worker FE	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓
Firm-Year Controls		✓		✓	✓
Worker Controls			✓	✓	✓
Occupation FE					✓
Observations	251934	251934	251934	251934	251934
Kleibergen-Paap rk Wald F statistic	533.395	253.765	319.210	174.680	127.925
Hansen J statistic	0.113	2.914	0.310	0.713	0.598
Hansen J statistic p-value	0.737	0.088	0.577	0.398	0.439

Notes: Second-stage IV estimations, in which World Import Demand with both contemporaneous and 10-year change in demand measures (see text) are used as excluded instruments for yearly net wages. Robust standard errors in parentheses with ***, **, and *respectively denoting significance at the 1%, 5% and 10% levels. Standard errors are clustered at the sector-year level. Firm-Year controls include log Assets, Log Apparent Labor Productivity, defined as total value added per worker, and log Employment. Worker controls include a skilled/unskilled dummy, the age of the worker and gender.

Appendix C: Difference-in-difference event study design: further results.

This section displays further results regarding (i) an augmented specification of the event study estimation including interactions between initial controls and year fixed effects and (ii) the other measures of import demand

(i) We first allow the specification to include the interaction between lagged controls and year fixed effects (as suggested by Goldsmith-Pinkham et al. 2020, see their section 5.2), with the benchmark WID. Results are plotted in Figure C1.

(ii) Second, Figure C2 plots results of the impact of various measures of WID on commuting distance, using the various weighting schemes (see the discussion about weights in paper, subsection 6.2).

Figure C1: Event study result (with initial controls x year dummies)

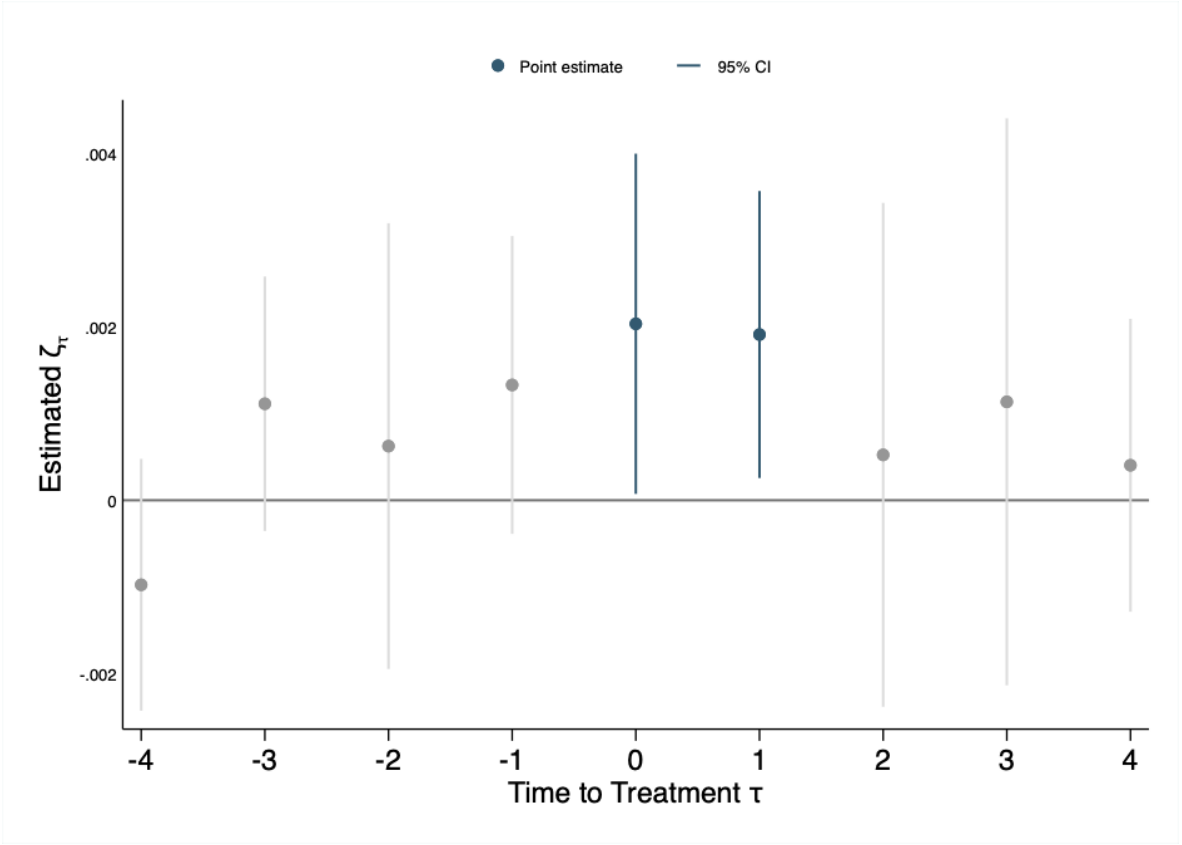


Figure C2: Event study results - Alternative WID measures.