Supplemental Appendix

"Racial and Ethnic Inequality and the China Shock"
Kahn, Oldenski, and Park

Census and American Community Survey Data

The primary datasets used in this paper are the 2000 5% U.S. Census and the American Community Surveys (ACS) for 2005 through 2019(US Census Bureau, 2005-2019a). We also use the 1980 and 1990 5% Censuses to measure pre-trends in employment gaps. We obtain data from the Census Integrated Public Use Micro Samples (US Census Bureau, 1960, 1970, 1980, 1990, 2000c). The Census and ACS samples include 5 and 1 percent of the U.S. population, respectively. We focus on 722 mainland commuting zones (CZs), which exclude those in Alaska and Hawaii, using the crosswalk from Public Use Microdata Areas (PUMAs) to CZs provided by Autor and Dorn (2013a).

We restrict attention to respondents aged 16 to 64 who do not reside in institutional group quarters. We classify observations as white if they report that they are not of Hispanic, Spanish, or Latino origin, and select "white" as their only race. We classify observations as Black if they are not Hispanic and select "Black" as any of their race choices (i.e., we categorize people who select multiple races as Black, as long as one of the races they select is). Finally, we categorize as Hispanic anyone who indicates that they are of Hispanic, Spanish, or Latino origin, regardless of race. For all of our analyses, we focus on just these three mutually exclusive (but not exhaustive) groups.

We aggregate observations to the CZ-race/ethnicity-year level using person weights. We define as employed anyone working in non-military employment. We define manufacturing jobs using the 1990 Census classification (taking values 100-392). The wage measure used in this paper is an hourly wage calculation. We replace top-coded annual wage and salary income with 1.5 times the top code value in that year. We define annual weeks worked using the categorical variable available in the Census and ACS datasets, imputing the midpoint of the category from 2000 (when both continuous and categorical variables are available) for all years. Hourly wages are top-coded adjusted annual income divided by the annual weeks

worked measure times usual hours worked per week and are missing if income, weeks, or hours are missing. We bottom-code wages to the first percentile in the national distribution each year and top code so that income for full-time, full-year work does not exceed the adjusted top-code value. Wages are inflation adjusted to the year 2012 using the Personal Consumption Expenditure Index (US Bureau of Economic Analysis, 1990-2019). We drop wage observations for the self-employed and military.

The main control variables (listed in footnote 14) are obtained from Autor, Dorn and Hanson (2021a) along with the Great Recession controls used in table 2, column 2. We use the 2000 Census to calculate the group-specific controls used in table 2, column 4. These include the share of the working age (16-64) population that graduated from college, the share of the working age female population that is employed, the share of employed that are working in a routine occupation and the average offshorability of the working age employed (Autor, Dorn and Hanson, 2013), and the share of the population that is 17 or less, 40-64, and 65+.

Our analyses aggregates several ACS waves (e.g., 2008-2013 or 2014-2019) and compares these to outcomes in the 2000 Census. When we aggregate, we first use sampling weights to aggregate within a year and then take an unweighted average across years.

Defining Import Exposure

To calculate the CZ-wide import penetration measure (equation 1) we follow Autor, Dorn and Hanson (2021a) (hereafter ADH). We use measures they provide for imports and domestic absorption at the 4-digit Standard Industry Classification (SIC) level. For a given 4-digit industry, we calculate the change in import exposure in year 2012 as the change in industry imports from 2012 compared to 2000 divided by domestic absorption. The latter is measured in 1991 and is equal to gross output plus imports minus exports.

We apply these changes in industry imports to the CZ-year level, following equation 1 in the text, i.e., summing across all industries weighting by the fraction of employment in the CZ in that industry in 2000. We use County Business Patterns (CBP) in 2000 from the U.S. Census Bureau to capture industry shares in the initial CZ employment.[1] CBP is an annual extension of the Census Bureau's economic censuses and provides employment in the private non-farm sector by county and 6-digit NAICS industry code. We use a version of these data provided by Acemoglu et al. (2016) who map to CZ-by-4-digit SIC cells.

---

[1]https://www.census.gov/programs-surveys/cbp/data.html

Our instrument for CZ-wide import exposure uses changes in Chinese imports from eight other high-income countries (Australia, Denmark, Finland, Germany, Japan, New Zealand, Spain, and Switzerland). Imports from these countries are also obtained from ADH. Domestic absorption is measured at a lag (1988 instead of 1991) and CZ-industry employment shares are also lagged, measured using the 1990 CBP.

For group-specific import exposure described in section III.C., we must use the U.S. Censuses to measure baseline employment shares by CZ, industry, and race/ethnicity (CBP data do not disaggregate by demographic group). We use the Census samples as described above to calculate employment shares from the 2000 Census (or 1990 Census for the instrumented version) at the CZ-industry-race/ethnicity level. Industries can only be measured at the 3-digit Census code level. We use the crosswalk of Autor, Dorn and Hanson (2019) to map 4-digit SIC codes to the 3-digit industry level (ind1990dd). Import exposure then sums the changes in imports from China across 3-digit industries (divided by 1991 domestic absorption aggregated to the 3-digit level in the same way), weighting by the fraction of employment in the CZ and subgroup in that 3-digit industry in 2000.

The instrument uses an analogous change in imports at the 3-digit industry level for the eight other high-income countries (divided by domestic absorption measured in 1988) and employment weights from the 1990 Census.

We have also explored a version of the CZ-wide measure that uses 3-digit Census industries and employed shares from the censuses, instead of 4-digit SIC industry codes and employment shares from CBP, and obtain similar results. These findings should allay concerns that our approach for measuring group-specific import exposure (which requires the higher level of industry aggregation) introduces too much error, and are available upon request.

To measure the China shares instrument used in column 3 of online appendix table OA.3, we obtain Chinese imports in the 8 other developed countries directly from the UN Comtrade Database at the 6-digit Harmonised System (HS) product-level (United Nations Statistics Division, 2000, 2012). We map these to 4-digit SIC industries using the crosswalk from Autor, Dorn and Hanson (2013). We then take the ratio of the sum of Chinese imports across the 8 countries for a given industry and the sum of all imports across the 8 countries for that industry. We take the difference across 2012 and 2000 and apply these to the CBP baseline employment shares in 1990.

For the NTR Gap shock, we obtain the spread between the NTR and non-NTR rate at the 8-digit HS product code level from Pierce and Schott (2016). We average this measure over

products within 4-digit SIC codes (using the same HS6 to 4-digit SIC crosswalk from above), weighting by product imports in 2000. We use County Business Patterns in 1990 to measure CZ employment shares across these industry categories.

Historical Outcomes, Figure 6

The top 2 panels are constructed using 5% Censuses from 1960, 1980, 1990, 2000, along with the 1970 1% state form 1 and 2 Censuses and 3-year averages of ACS waves from 2001-2019. We obtain data from the Census Integrated Public Use Micro Samples (US Census Bureau, 1960, 1970, 1980, 1990, 2000c). We restrict to ages 16-64 not living in institutional group quarters. We exclude the military from employment. We use the most detailed definition of race available in the given year.

The manufacturing weekly wage premium is calculated as follows. We exclude the self employed and replace top coded values for annual wage and salary income with 1.5 times the top code amount. We use the categorical weeks worked variable because not all years include the continuous weeks worked variable. We impute the midpoint for the category using 1980 continuous values for 1960, 1970, and 1980, 1990 continuous values for 1990, and 2000 continuous values for all years from 2000 onwards. We calculate weekly wages by dividing annual income by weeks worked. We then bottom code the first percentile in the national distribution for that year and top code so that weekly wages times full-year work (50 weeks) does not exceed the top coded value. We inflation adjust to 2012 using the PCE.

To obtain manufacturing wage premia, we then regress log weekly wages on a manufacturing indicator and interactions with Black and Hispanic. We control for group main effects, state fixed effects, and age-sex-education cell fixed effects. Age groups are decades and education groups are the general educ categories. We estimate these regressions separately by year and use sampling weights.

The bottom left panel on union membership uses CPS outgoing rotation groups (ORG) from 1983-2019 and March supplements from 1990-2019. These are also obtained from IPUMS (US Census Bureau, 1983-2019b). We restrict to ages 16-64. We use as much detail as available in a given survey year to classify race. We plot group differences in union membership though results are similar if we also include those covered by a union but who are not themselves members.

Job-to-Job Flows

For table 3, we derive race and ethnicity-specific quarterly job-to-job flows in year 2000 from the Job-to-Job Flows (J2J) Explorer, which is based on Longitudinal Employer-Household Dynamics (LEHD) data (US Census Bureau, 2000a). J2J provides a set of statistics on job mobility, such as the number of job-to-job transitions between 3-digit NAICS industries and hires and separations to and from employment. We aggregate the industry-level transitions up to the manufacturing and non-manufacturing sectors and take the average of the quarterly transitions in the third and fourth quarters of 2000 because the J2J series started in the third quarter of 2000. J2J do not provide the statistics at the subgroup level for many states because they do not meet U.S. Census Bureau publication standards. So, we aggregate the values only for the available states: Alaska, California, Delaware, Florida, Hawaii, Idaho, Illinois, Iowa, Kansas, Minnesota, Missouri, Nebraska, New Jersey, New York, North Carolina, Oklahoma, Oregon, Pennsylvania, South Carolina, South Dakota, Texas, Washington, and Wisconsin. To calculate the job-to-job flow rates and separation rates, we divide the job-to-job transitions and separations by total employment in the sectors from the Quarterly Workforce Indicators (QWI) for the same period and the same states (US Census Bureau, 2000b). The QWI is also based on LEHD, so it should be consistent with J2J.

References: Online Appendix

Acemoglu, Daron, David Autor, David Dorn, Gordon H. Hanson, and Brendan Price. 2016. "Data for Import Competition and the Great US Employment Sag of the 2000s." retrieved from https://ddorn.net/data.htm.

Autor, David, David Dorn, and Gordon Hanson. 2013. "The China Syndrome: Local Labor Market Effects of Import Competition in the United States." *American Economic Review*, 103(6): 2121–2168.

Autor, David, David Dorn, and Gordon Hanson. 2019. "When Work Disappears: Manufacturing Decline and the Falling Marriage Market Value of Young Men." *American Economic Review: Insights*, 1(2): 161–78.

Autor, David, David Dorn, and Gordon Hanson. 2021a. "Data for On the Persistence of the China Shock." retrieved from https://ddorn.net/data.htm.

Autor, David, David Dorn, and Gordon Hanson. 2021*b*. "On the Persistence of the China Shock." *Brookings Papers on Economic Activity*, 2021(Fall): 448–456.

Autor, David H., and David Dorn. 2013*a*. "Data for The Growth of Low-Skill Service Jobs and the Polarization of the US Labor Market." retrieved from https://ddorn.net/data.htm.

Autor, David H., and David Dorn. 2013*b*. "The Growth of Low-Skill Service Jobs and the Polarization of the US Labor Market." *American Economic Review*, 103(5): 1553–97.

Pierce, Justin R., and Peter K. Schott. 2016. "Data for The Surprisingly Swift Decline of US Manufacturing Employment." retrieved from https://sompks4.github.io.

United Nations Statistics Division. 2000, 2012. "UN Comtrade Database." retrieved from https://comtrade.un.org/.

US Bureau of Economic Analysis. 1990-2019. "Personal Consumption Expenditures (PCE)." retrieved from https://fred.stlouisfed.org/series/PCE.

US Census Bureau. 1960, 1970, 1980, 1990, 2000*c*. "US Census: Census Integrated Public Use Micro Samples (IPUMS)." retrieved from https://www.ipums.org/.

US Census Bureau. 1983-2019*b*. "Current Population Survey (CPS): Census Integrated Public Use Micro Samples (IPUMS)." retrieved from https://www.ipums.org/.

US Census Bureau. 2000*a*. "Longitudinal Employer-Household Dynamics (LEHD) Job-to-Job Explorer." retrieved from https://j2jexplorer.ces.census.gov/.

US Census Bureau. 2000*b*. "Quarterly Workforce Indicators (QWI) Explorer." retrieved from https://qwiexplorer.ces.census.gov/.

US Census Bureau. 2005-2019*a*. "American Community Survey (ACS): Census Integrated Public Use Micro Samples (IPUMS)." retrieved from https://www.ipums.org/.

Additional Tables and Figures

Figure OA.1: Minority-White Differences by Occupation
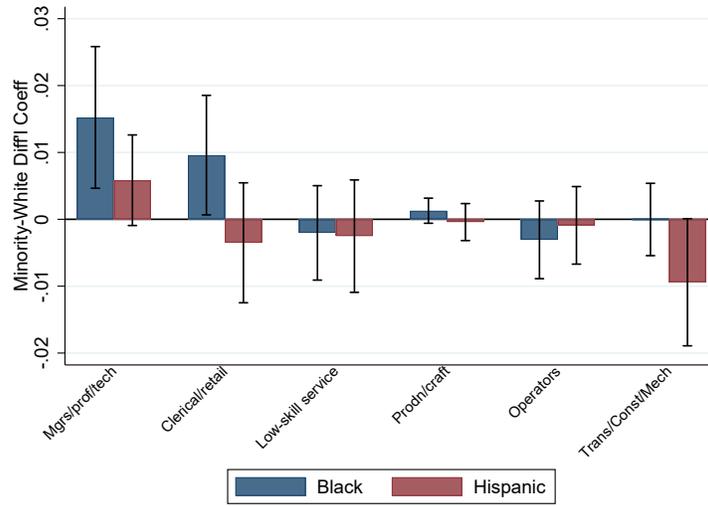
Notes: We report minority-white differential impacts of CZ-wide import exposure on occupation-specific employment per adult population within the race/ethnic group. Occupation categories come from the Level 1 Autor and Dorn (2013*b*) classification. See figure 4.

Figure OA.2: Distributions of Group-Specific Measures of Import Exposure

White Mean: 1.09, Black Mean: .87, Hispanic Mean: 1.27
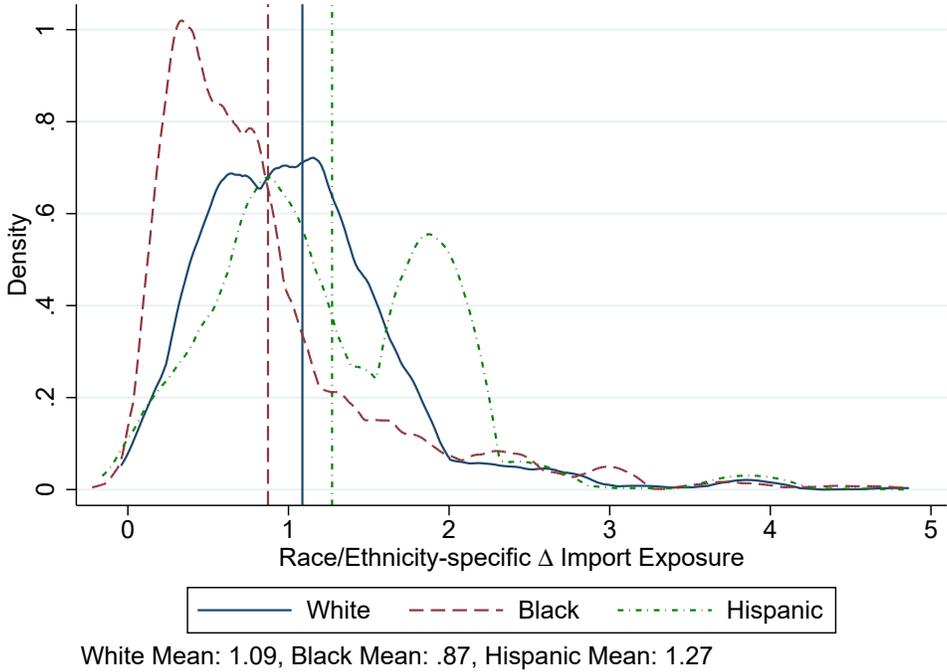
Notes: We plot the distributions across CZs of group-specific change in import exposure from China (IP) from 2000-2012. Relative to the import exposure measure defined in equation 1 and taken from Autor, Dorn and Hanson (2021*b*), the measure plotted here allows for different baseline industry employment shares by race/ethnicity as measured in the 2000 Census ($\Delta IP_{rct} = \sum_i \frac{Emp_{irc}}{Emp_{rc}} \frac{\Delta M_{it}}{Norm_i}$). White, Black, and Hispanic populations are mutually exclusive (but not exhaustive). Densities are weighted by race/ethnicity populations in 2000. Group-specific means are indicated with vertical lines. For clarity, the density plots (but not the mean lines) omit 2 outlier CZs with exposures greater than 9.

Figure OA.3: Minority-White Baseline Employment Shares by Industry
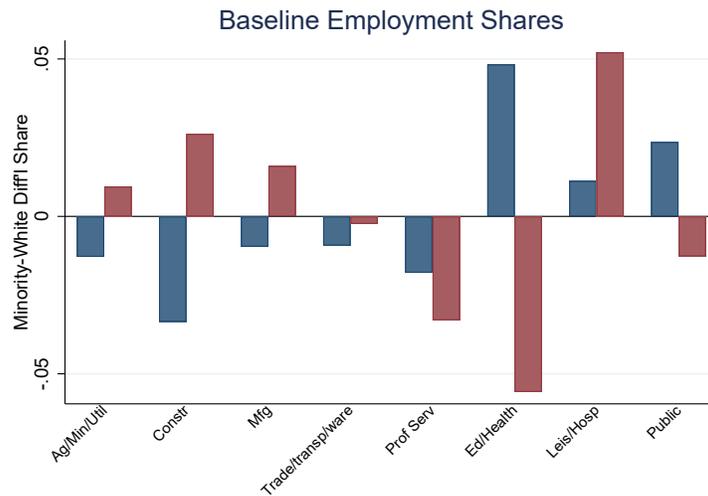
Notes: We report minority-white differentials in the share of employment that is in a given industry in 2000. Industries are defined with the following NAICS codes: Ag/Min/Util (11, 21, 22); Constr (23); Mfg (31-33); Trade/transp/ware (42, 44-45, 48-49); Prof Serv (51, 52, 53, 54, 55, 56); Ed/Health (61, 62); Leis/Hops (71, 72, 81); Public(92).

Table OA.1: Summary Statistics

	White			Black			Hispanic		
	2000	2010	2016	2000	2010	2016	2000	2010	2016
ΔIP (ADH)	1.10			1.01			1.06		
	(0.75)			(0.73)			(0.64)		
Mfg Emp per pop	11.03			8.28			9.60		
	(4.55)			(4.89)			(5.13)		
Non-mfg Emp per pop	62.56			52.30			49.81		
	(5.58)			(7.11)			(5.93)		
Overall Emp per pop	73.59			60.58			59.42		
	(4.47)			(4.82)			(4.59)		
Log Hourly Wage	2.82			2.65			2.54		
	(0.17)			(0.15)			(0.09)		
Changes from 2000:									
log Mfg Emp		-0.30	-0.32		-0.44	-0.42		-0.32	-0.39
		(0.11)	(0.12)		(0.19)	(0.21)		(0.17)	(0.19)
log Non-Mfg Emp		-0.01	0.03		0.03	0.11		0.16	0.22
		(0.03)	(0.03)		(0.07)	(0.06)		(0.07)	(0.08)
log Overall Emp		-0.05	-0.02		-0.02	0.05		0.09	0.14
		(0.03)	(0.03)		(0.07)	(0.05)		(0.06)	(0.06)
log Hourly Wage		0.02	0.06		-0.04	-0.03		-0.00	0.05
		(0.04)	(0.05)		(0.05)	(0.06)		(0.04)	(0.04)
# CZs	722	722	722	722	722	722	722	722	722

Notes: We summarize CZ cells by race/ethnicity, weighting by group-specific population in the CZ in 2000. 2000 data are from the Census. 2010 reflects an unweighted average of 2008-13 American Community Survey (ACS) waves. 2016 reflects an unweighted average of 2014-19 ACS waves. Employment variables are per adult (age 16-64) non-institutionalized group-specific population. Changes are in log employment per population from 2000. Log hourly wages are annual wage and salary income divided by annual weekly hours time usual hours per week, adjusted to 2012 dollars using the PCE price index, and exclude self-employed and military. All employment measures exclude military employment.

Table OA.2: First-Stage Regressions

	(1)	(2)	(3)
Panel A:		ADH Instrument	
Dependent Variable:		ΔIP (ADH)	
ΔIP IV	0.426	0.470	0.504
	(0.073)	(0.073)	(0.060)
Observations	722	722	722
R-squared	0.653	0.674	0.804
F-stat on instrument	34	42	71
Panel B:		Chinese Import Share Instrument	
Dependent Variable:		ΔIP (ADH)	
Chinese Shares IV	0.247	0.292	0.267
	(0.033)	(0.035)	(0.025)
Observations	722	722	722
R-squared	0.670	0.710	0.814
F-stat on instrument	56	69	119
Panel C:		NTR Gap Instrument	
Dependent Variable:		ΔIP (ADH)	
NTR GAP IV	7.726	8.362	10.473
	(2.302)	(2.797)	(3.399)
Observations	722	722	722
R-squared	0.565	0.596	0.717
F-stat on instrument	11	9	9
Panel D:		Own-Group Shock	
Dependent Variable:		Group-Specific ΔIP	
Group-specific IV	1.005	0.742	0.610
	(0.111)	(0.101)	(0.070)
Observations	722	722	722
R-squared	0.766	0.712	0.713
F-stat on instrument	83	54	76
White	X		
Black		X	
Hispanic			X

Notes: See table 1. We regress the indicated import exposure measure in the contemporaneous year minus that in 2000 on the import exposure instruments, separately for white, Black, and Hispanic, including full controls. Panel A uses as an instrument changes in imports from China for other developed countries applied to lagged (race-specific or CZ-wide) employment shares. Standard errors are clustered on state. Kleibergen-Paap Wald F-stats are reported. Models are weighted by race-specific CZ working-age population in 2000.

Table OA.3: Impacts of Import Exposure: Alternative Identification Strategies

Time Period:	2010-2000				2016-2000			
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Dependent Variable:				Δ log Mfg Emp Rate				
ΔIP (ADH)	-0.069	-0.029	-0.087	-0.143	-0.095	-0.029	-0.107	-0.124
	(0.023)	(0.009)	(0.020)	(0.051)	(0.027)	(0.011)	(0.027)	(0.051)
$\Delta IP * Black$	0.048	0.021	0.028	0.049	-0.003	-0.017	-0.003	0.018
	(0.044)	(0.016)	(0.035)	(0.068)	(0.048)	(0.019)	(0.034)	(0.084)
$\Delta IP * Hispanic$	-0.001	-0.027	0.026	-0.035	-0.007	-0.049	0.016	-0.093
	(0.038)	(0.016)	(0.029)	(0.065)	(0.036)	(0.017)	(0.035)	(0.070)
Observations	2,052	2,052	2,052	2,052	2,059	2,059	2,059	2,059
Dependent Variable:				Δ log Non-Mfg Emp Rate				
ΔIP (ADH)	0.002	0.004	0.004	0.003	0.003	0.005	0.005	0.011
	(0.004)	(0.002)	(0.004)	(0.010)	(0.004)	(0.002)	(0.004)	(0.009)
$\Delta IP * Black$	0.029	0.012	0.023	0.036	0.029	0.009	0.021	0.057
	(0.011)	(0.006)	(0.013)	(0.026)	(0.011)	(0.005)	(0.011)	(0.034)
$\Delta IP * Hispanic$	-0.023	-0.010	-0.023	0.027	-0.023	-0.004	-0.025	0.011
	(0.011)	(0.010)	(0.014)	(0.026)	(0.012)	(0.012)	(0.016)	(0.032)
Observations	2,166	2,166	2,166	2,166	2,166	2,166	2,166	2,166
Dependent Variable:				Δ log Overall Emp Rate				
ΔIP (ADH)	-0.009	-0.001	-0.012	-0.033	-0.013	-0.001	-0.014	-0.027
	(0.006)	(0.002)	(0.004)	(0.013)	(0.005)	(0.002)	(0.004)	(0.010)
$\Delta IP * Black$	0.028	0.012	0.023	0.023	0.019	0.005	0.015	0.031
	(0.011)	(0.005)	(0.010)	(0.024)	(0.008)	(0.003)	(0.007)	(0.017)
$\Delta IP * Hispanic$	-0.011	-0.012	-0.006	0.017	-0.008	-0.010	-0.006	0.002
	(0.007)	(0.005)	(0.007)	(0.017)	(0.008)	(0.006)	(0.008)	(0.017)
Observations	2,166	2,166	2,166	2,166	2,166	2,166	2,166	2,166
Dependent Variable:				Δ log Hourly Wages				
ΔIP (ADH)	-0.005	-0.000	-0.008	-0.024	-0.009	0.001	-0.011	-0.018
	(0.008)	(0.003)	(0.007)	(0.013)	(0.008)	(0.003)	(0.007)	(0.015)
$\Delta IP * Black$	0.029	0.003	0.021	0.024	0.035	0.004	0.022	0.020
	(0.010)	(0.003)	(0.008)	(0.020)	(0.009)	(0.003)	(0.011)	(0.027)
$\Delta IP * Hispanic$	0.021	0.003	0.016	0.064	0.017	0.003	0.011	0.057
	(0.009)	(0.006)	(0.010)	(0.024)	(0.006)	(0.004)	(0.006)	(0.022)
Observations	2,166	2,166	2,166	2,166	2,166	2,166	2,166	2,166
Main IV	X				X			
OLS		X				X		
Shares IV			X				X	
NTR IV				X				X

Notes: Column 1 reproduces the column 1 specification from table 1. Column 2 estimates OLS regressions. Column 3 instruments with changes in Chinese market shares, rather than levels of imports, in other high income countries. Column 4 instruments for the shock and its group interactions with the NTR Gap IV and its group interactions. Standard errors in parentheses clustered by state.