Comment on Jonathan Rothwell's Critique of Autor, Dorn, and Hanson (2013)

David Autor, MIT and NBER

David Dorn, University of Zurich and CEPR

Gordon Hanson, UC San Diego and NBER

March 3, 2017

In a series of recent papers, Rothwell (2016a, 2016b, 2017) claims that the empirical results in Autor, Dorn, and Hanson (2013, ADH hereafter) on the impact of import competition with China on local labor markets in the United States are subject to specification bias, not robust to separating the sample into sub-periods, and subject to other weaknesses in analytical approach. In this note, we respond to the main thrust of Rothwell's comments on ADH. To provide background on the origins or our response, after we provided comments on Rothwell (2016a), which identified fundamental flaws in the empirical analysis, Rothwell (2016b) was submitted to an academic journal. This note contains much of the body of our referee report on that submission (which was rejected for publication as a comment on ADH). Although Rothwell (2017) contains minor extensions on Rothwell (2016b), our sense is that the main thrust of the critique is the same. For the purpose of clarifying our stance on the series of papers that Rothwell has produced, we thought it would be instructive to provide our formal (and heretofore private) response to his critique on our scholarly work.

It is helpful to begin by briefly summarizing the results in Autor, Dorn, and Hanson (2013). ADH analyze the impact of import competition from China on local labor markets in the U.S. They use the 722 commuting zones (CZs) in the continental U.S. to represent local labor markets and examine the impact of the China trade shock over two time periods, 1990 to 2000 and 2000 to 2007. They measure the trade shock using the growth in China imports per worker (see their equation 3) and instrument for this value using growth in imports from China in other high-income countries (see their equation 4). ADH estimate the impact of the China trade shock on changes in CZ manufacturing employment; the CZ working-age population; CZ non-manufacturing employment, unemployment, population not in the labor force, and population receiving Social Security Disability Insurance (SSDI); CZ average log weekly wages; CZ government transfer receipts per capita; and CZ household income by source. They find that CZs more exposed to import competition from China experience significantly larger reductions in manufacturing employment, increases in unemployment and non-participation in the labor force, decreases in wages in the non-manufacturing sector, increases in government transfer receipts, and reductions in household income. They do not find significant impacts of trade shocks on changes in CZ non-manufacturing employment, working-age populations, or manufacturing wages.

Main Critique: The ADH regression specification introduces bias into the estimation of the impact of rising import exposure on changes in local labor market outcomes. When corrections are introduced, some results in ADH are weakened.

To begin, it is useful to review the primary regression specification in ADH, which is

$$\Delta y_{it} = \gamma_t + \beta_1 \Delta IPW_{uit} + \lambda_1 M_{it} + X'_{it}\delta + e_{ct}, \tag{1}$$

where Δy_{it} is the decadal change in labor-market outcome y in commuting zone i, the change in import exposure ΔIPW_{uit} is the key explanatory variable of interest, and M_{it} is the lagged value of the initial manufacturing employment share (which is the the control variable on which Rothwell (2016b) focuses). When estimating this model for the interval between 1990 and 2007, ADH stack the 10-year equivalent first differences for two periods, 1990 to 2000 and 2000 to 2007, and include separate time dummies for each decade (in the vector γ_t). The change in import exposure ΔIPW_{uit} is instrumented by the variable ΔIPW_{oit} as described in ADH equation (4). This stacked first-difference model is a three-period fixed effects model but with less restrictive assumptions on the error term (see ADH footnote 26). The vector X_{it} contains dummies for Census geographic divisions and controls for CZs' start-of-decade share of the working-age population that is college-educated, share that is foreign born, share of women in total employment, share of employment in routine occupations, and average offshorability index (see notes to Table 3 in ADH). These controls allow changes in outcomes to be a function of initial conditions, time trends to vary by geographic region, and the aggregate time trend to vary by decade. Standard errors are clustered at the state level.

The main critique in Rothwell (2016b) is that some ADH coefficient estimates fall in magnitude and lose statistical significance when one includes in the regression the interaction between a time dummy for the 2000s and the initial-period CZ share of employment in manufacturing M_{it}. Rothwell (2016b) argues that this interaction captures unspecified adverse shocks to manufacturing in the 2000s, which ADH confound with the China trade shock.

Observation 1: The main methodological critique provided does not follow from its own logic.

Rothwell begins with the observation that the two time periods ADH consider—the 1990s and the 2000s—differ in macroeconomic conditions and in terms of the magnitude of the increase import competition from China. Rothwell (2016b) reports on page 8 [Start of quotation]:

1. Outcomes tend to be worse in the second period (2000 to 2007)

2. Import exposure tends to be higher in the second period

3. CZs that are more import exposed have worse outcomes in the second period relative to the first

[Quotation continues] ADH data show that import exposure is over twice as high during the second period for the average CZ. Examining 18 core outcomes—16 of which are directly analyzed by ADH—I show in Table 1 that all of these outcomes are worse in the second period—even for CZs in the bottom quartile of import-exposure—but this is particularly true for highly exposed CZs—those in the top quartile of import exposure. In summary, the time-trend variable both predicts worse outcomes generally but especially for the most exposed CZs, *complicating the stacked model framework* [End of quotation, italics added].

Nothing in this chain of logic "complicates" inference in the stacked model framework; in fact, it is fully consistent with a stacked model. Using Eqn. (1) to formalize these three claims, one would write: (1) outcomes tend to be worse in the second period, meaning that Δy_{it} is more negative in the second period; (2) import exposure tends to be higher in the second period, meaning that ΔIPW_{uit} is larger in the second period; and (3) CZs that are more import exposed have worse outcomes in the second period relative to the first, implying that β_1 is plausibly the **same** in both periods. That is, with β_1 constant across periods, a larger rise in the independent variable forces a larger fall in the dependent variable, which is precisely what is implied by the ADH results *and* by the three observations above. A stacked model is thus entirely in line with observed changes in manufacturing employment and trade exposure over the two periods considered in the ADH analysis.

Observation 2: Rothwell's (2016b) main critique is not formally tested. When the appropriate test is conducted, there is strong support for the findings in ADH.

Rothwell (2016b) proposes a solution to time variation in trade shocks and macroeconomic conditions between the two decades, which is to add to the regression a single interaction term, between the initial manufacturing share and a time dummy. This approach is arbitrary and selective. If the issue is that the impact of the trade shock in non-constant over time, then the obvious solution is to relax the parameter homogeneity assumption on the treatment effect. Rothwell (2016b), however, never performs this exercise. He chooses instead to relax parameter homogeneity for a single control variable, an approach which does not follow from the logic of the critique.

To implement the logic of Rothwell (2016b), we estimate three variations on the above ADH specification for a wide range of outcomes. Consider the following generalized version of Eqn. 1,

$$\Delta y_{it} = \gamma_t + \beta_1 \Delta IPW_{uit} + \beta_2 \left(\Delta IPW_{uit} \times [t \geq 2000]\right) + \lambda_1 M_{it} + \lambda_2 \left(M_{it} \times [t \geq 2000]\right) + X'_{it}\delta + e_{ct}, \quad (2)$$

Table 1 shows estimation results for this exercise. First, we estimate the baseline ADH specification, which is shown in column 1 of Table 1, as well as in column 1 of Table 3 in Rothwell (2016b) (where we include ADH outcomes that he ignores). In Eqn. (2), this specification involves setting β_2 and λ_2 equal to zero. Second, we estimate Rothwell's specification, which adds to the ADH model an interaction between the initial manufacturing share and a time dummy for the 2000s, as shown in column 2 of Table 1 as well as in column 2 of Table 3 in Rothwell (2016b). In Eqn. (2), this specification sets β_2 equal to zero but relaxes the constraint on λ_2. Third, we report the specification that follows from the logic of Rothwell, which adds to the ADH model two interaction terms, one for the trade shock and a time dummy and another for the initial manufacturing employment share and a time dummy, as shown in column 3 of Table 3. In Eqn. (2), this specification involves no constraints on β_2 or λ_2. Because we report results for 22 outcome variables in three specifications, two of which include time interactions for key regressors, we simplify the presentation by summarizing parameter estimates in terms of their sign and statistical significance. The color coding in Table 1 indicates whether the coefficient estimate for β_1 in columns 1 and 2 has a positive sign (green) or a negative sign (orange). The color coding in column 3 indicates the sign of the average of the period-specific effects of β_1 for the first period and $\beta_1 + \beta_2$ for the second period.

When comparing column 1 (ADH) to column 2 (Rothwell, 2016b), two findings are apparent. First, *the main ADH effects on manufacturing employment are strongly robust.* Whether the outcome variable is the change in manufacturing workers as a share of the working-age population, the change in manufacturing workers as a share of the college population, the change in manufacturing workers as a share of the non-college population, or the log change manufacturing employment, the impact of the China trade shock is negative and very precisely estimated. Second, we see that for 12 outcome variables the impact coefficient on the China trade shock loses statistical significance in going from column 1 to column 2 (although coefficient signs are the same in 18 of the 22 regressions). Seven of the 12 cases in which the trade shock loses significance from column 1 to column 2 have an outcome variable that is a wage or an income measure. ADH describe at length the difficulty in using observed changes in average earnings to make inferences about the impact of trade on wages. We discuss this issue in more detail below. For now, however, we simply note in going from column 1 to column 2 it appears that the impact of the trade shock on some ADH outcomes loses significance.

As we have stated, the column 2 regression, which represents the specification favored by Rothwell (2016b), has no clear rationale. If the issue is time variability in the impact of trade on outcomes of interest, then one should allow the impact coefficient to vary over time. We implement this approach in column 3 of Table 1. Comparing columns 1 and 3, we see that coefficient signs on trade impacts are the same for all outcome variables. Thus, allowing for parameter heterogeneity in the trade shock in no way changes the qualitative impact of trade on the outcomes that ADH examine. Turning to statistical significance, for each regression column 1 summarizes the results of a t-test on $\beta_1 = 0$ (i.e., the null is no trade impact on average over the two time periods), under the maintained assumptions that $\beta_2 = 0$ and $\lambda_2 = 0$, and column 3 summarizes the results of an F-test on the joint hypothesis that $\beta_1 = 0$ and $\beta_2 = 0$ (i.e., the null is no trade impact in either time period). Of the 19 outcomes for which we reject the null of no trade impact in column 1, we reject the null of no trade impact for 15 of these cases in column 3. Thus, for most of the outcomes that ADH consider we reject the null of no trade impact either under the restriction of parameter homogeneity across time or under parameter heterogeneity across time. *When we apply the statistical test implied by the logic of Rothwell (2016b), we see no support for his claim that parameter restrictions bias the results* *in favor of finding trade impacts.*

It is worthy of note that the changes in significance between columns 1 and 3 do not move monotonically. Although for one outcome (the percent change in average household total income) the trade shock drops from a 1% level of significance in column 1 to a 10% level in column 3, for three other outcomes (the change in the share of non-manufacturing workers in the population, the log change in non-manufacturing employment, the log weekly wage in non-manufacturing) coefficients move from insignificant or marginally significant in column 1 to strongly significant in column 3.

To summarize, when we follow the logic of the critique in Rothwell (2016b) to permit time variability in the treatment effect, we see that the ADH results on the impact of trade shocks on a wide range of labor-market outcomes are strongly robust. The exceptions to this pattern are primarily for outcomes related to log average weekly wages, which we discuss next.

Observation 3: Most of the results affected by the additional control relate to changes in earnings, which ADH acknowledge are poorly measured in their data.

Three of the four outcomes for which the trade shock loses statistical significance in going from column 1 to column 3 of Table 1 relate to log weekly wages. ADH make clear that the data they use in their analysis are not well suited to study changes in wages (see their section IV.B). Because they do not have longitudinal data on individual earnings, they cannot say whether observed changes in earnings in their data are due to changes in wages for individual workers or to differential selection out of employment by high or low-wage workers. This potential for compositional bias complicates estimating the impact of trade shocks on changes in wages using repeated cross-section data.

As seen in Table 1, the sign and the magnitude of the wage impacts of trade shocks appear to be inconclusive. If there were no other results in the literature on how trade with China affects earnings for U.S. manufacturing workers, we would have to leave the matter here. However, Autor, Dorn, Hanson, and Song (2014) perform the exact type of longitudinal analysis that can address the measurement problems that plague observed average wages in manufacturing. They find that controlling for an extensive set of worker and employer characteristics in the pre-sample period, workers initially more exposed to the China trade shock have lower annual earnings, lower cumulative income over the 1991 to 2007 period, and greater uptake of SSDI. If one wants to understand the impact of trade with China on earnings for U.S. manufacturing workers, one should turn to these results and not to those in ADH.

Observation 4: Other comments.

Rothwell (2016b) also makes several points which are not germane to ADH, but which are worth mentioning briefly. One is that the China trade shock is driven by a relatively small number of manufacturing industries (section 2.C). We do not see how this counts as a critique of ADH. The impact of the China trade shock is identifiable precisely because the country's comparative advantage is concentrated in manufacturing and, within manufacturing, in a limited set of labor-intensive products. The combination of concentrated comparative advantage and accelerated economic growth is what made trade with China so disruptive for particular U.S. industries and regions. A second point that is not germane to ADH is that the China trade shock is not correlated with changes in life expectancy over the extended time period of 1990 to 2011 or with levels of household income and income inequality in 2011 to 2014 (section 3.C). Life expectancy, which ADH do not examine, is problematic both because of the slow rate at which it changes and because data on the variable are missing for many commuting zones. The analysis of how trade shocks over 1990 to 2011 affect outcomes in *levels* in 2011 to 2014 uses a specification at odds with the first-difference regression in ADH, which controls for CZ fixed effects. The regressions in Rothwell (2016b) that use levels for outcomes lack such controls and thus are difficult to compare to ADH.

Summary

Rothwell (2016a, 2016b, 2017) claims that some results in ADH are sensitive to the inclusion of additional controls for period-specific manufacturing shocks. However, further investigation reveals that (a) the ADH main results on manufacturing employment are robust to these and other controls, (b) the main methodological critique provided does not follow from Rothwell's own logic, (c) when we follow the logic of the critique and allow for time-varying impacts of trade shocks, we reject the null of no trade impacts in 17 of the 22 ADH outcomes considered, and (d) most of the outcomes for which the impact of trade shocks are sensitive to the specification relate to difficult-to-measure average wages. We conclude that the main methodological critique provided by Rothwell is unfounded. The critique he provides does not follow from his own logic, the modification applied to the ADH specification does not actually test for the parameter heterogeneity that Rothwell claims is present, and when the specification is modified to test that claim the ADH findings are strongly supported.

References

[1] Autor, David H., David Dorn, Gordon H. Hanson. 2013. "The China syndrome: local labor market effects of import competition in the United States." *American Economic Review*, 103(6):2121–2168.

[2] Autor, David H., David Dorn, Gordon H. Hanson, Jae Song. 2014. "Trade adjustment: worker level evidence." *Quarterly Journal of Economics*, 129(4):1799–1860.

[3] Rothwell, Jonathan. 2016a. "Trade competition and local economic performance: a reanalysis of recent work." Mimeo, Gallup.

[4] Rothwell, Jonathan. 2016b. "Trade competition and local economic performance: a comment on recent work." Mimeo, Gallup.

[5] Rothwell, Jonathan. 2017. "Cutting the losses: reassessing the cost of import competition to workers and communities." Mimeo, Gallup.

Table 1: Autor, Dorn and Hanson (AER 2013): Robustness of Regression Estimates to Alternative Specifications. Stacked First Difference Models, 1990-2000 and 2000-2007, 2SLS Estimates.

	Regression Specification			
	ADH Baseline Specification	ADH plus Mfg*Time (Comment Table 3)	ADH plus Mfg*Time and IPW*Time	
Outcome Variables:	Significance ipw	Significance ipw	Joint Sign. ipw, ipw*t	Significance ipw*t
Change in	(1)	(2)	(3)	(4)
1 mfg emp/pop	***	***	***	
2 n-mfg emp/pop			***	***
3 unemp/pop	***		**	
4 nilf/pop	***		***	***
5 mfg emp/pop, coll	***	***	**	
6 unemp/pop, coll	***		**	
7 mfg emp/pop, n-coll	***	***	***	
8 unemp/pop, n-coll	***			
9 log mfg emp	***	***	***	
10 log nmfg emp			*	**
11 log unemp	***	**	**	
12 log nilf	*		***	***
13 log weekly wage	***			
14 log weekly wage coll	**			*
15 log weekly wage n-coll	***			*
16 log weekly wage mfg		**	**	
17 log weekly wage nmfg	***		**	**
18 log transfers pc	***	**	***	**
19 avg hh tot income (%chg)	***		*	
20 avg hh wage inc (%chg)	***		***	**
21 med hh tot income (%chg)	***		***	**
22 med hh wage inc (%chg)	***	*	***	***

Notes: N=1444 (722 commuting zones x 2 time periods). Column (1) reports results the significance pattern of the imports-per-worker (ipw) variable as reported in ADH (2013), based on stacked first differences regressions with full controls (*, **, *** indicate rejection of the null hypothesis of a zero coefficient at the 10%, 5%, 1% significance level). Column (2) reports the significance pattern based on regressions that add an interaction between the manufacturing share and the period dummy to the setup of column (1), corresponding to the Table 3 regressions in the comment. Column (3) adds to the column (2) model and interaction between the imports-per-worker variable and the period dummy, and reports the outcome of a hypothesis test for the joint signficiance of the ipw variable and its time interaction. The color shading in columns (1) and (2) indicates whether the coefficient of the trade shock is positive (green) or negative (red), while the color shading in column (3) corresponds to the coefficients of the trade shock averaged over both periods. Column (4) indicates the significance of the interaction term between ipw and the period dummy.