A Reply to Campbell and Mau

Nick Bloom, Mirko Draca and John Van Reenen, November 20th 2020

A shortened version of this response is forthcoming in the Review of Economic Studies.

In Bloom, Draca and Van Reenen (2016, "BDVR") we have a set of nine results on the impact of Chinese trade. The first three showed that Chinese trade increased technical change in European firms measured by patents, productivity and IT adoption. The last six showed that Chinese trade led to reallocation towards more technologically advanced firms: those with more patents, higher productivity and IT adoption had faster growth and lower exit rates. Campbell and Mau (2020, "CM") argue that the effects of Chinese imports on patenting are sensitive to specification changes.

The precise functional form of the dependent variable in long differenced models does not drive our results

In their Table 1, CM argue that the percentage fall in patents is no different in the "treatment" sectors (where Chinese quotas were removed) compared to the "control" sectors (where quotas were not removed). This occurs when they perform a sector level analysis and drop half the data we used in our original paper (years 1999-2004) and use only the other half (years 2000-2005). It is unclear why they do this as both years span China joining the WTO in 2001. Indeed, given the earlier years are less influenced by censoring (the lag between patent applications and grant date) these earlier years would seem preferable to use. Table 1 below shows that the earlier years and the sample as a whole are consistent with our main results of a positive treatment effect.

In Table 2 we show two standard alternative transformations for dealing with patent counts of zero – the arc-percentage change (Davis et al., 1996) and the Inverse Hyperbolic Sine (Burbridge et al., 1988). These give similar results to the log difference we used in our baseline. Hence, the functional form of the dependent variable does not lead to fundamentally different results in long-differenced models.

Count data models show a positive impact of Chinese imports

CM point to coding errors in our original Table 7. Column (1) of Table 3 below reproduces our original result, and column (2) corrects for the coding errors (equivalent to CM Table 2, columns (1) and (3) respectively). However, CM's column (4) omits the industry dummies that we use to control for sector heterogeneity. Our baseline *long differenced* regressions in Tables 1-5 of BDVR removes these industry fixed-effects through differencing, but they are necessary in the *levels* count data models (e.g., due to variations in intensity to file patents).

A second issue with the column (2) specification in our Table 3 is that it does not control for the initial conditions for Chinese imports. To see why this is potentially important, consider the model:

$$PAT_{ijkt} = \exp\left(\alpha IMP_{jkt}^{CH} + f_{kt} + \eta_i\right) V_{ijkt} \quad (1)$$

where PAT_{ijkt} is the count of patents of firm *i* in industry *j* in county *k* at time *t*, IMP_{jkt}^{CH} is the firm's exposure to Chinese imports, f_{kt} are country by time dummies, η_i is a firm fixed effect and V_{ijkt} an idiosyncratic error term. We can approximate η_i by a linear function of industry dummies ($SIC4_j$), the initial patent stock, $\overline{PAT}_{ijk0}$, and initial Chinese imports, $\overline{IMP}_{jk0}^{CH}$. Formally, the assumption is:

$$\exp(\eta_i) = \exp(SIC4_j + a_1 \overline{PAT}_{ijk0} + a_2 \overline{IMP}_{jk0}^{CH}) U_i \, , \quad (2)$$

where U_i has mean 1 and is independent of all conditioning variables. Thus, the equation we take to the data is:

$$E(PAT_{ijkt} | conditioning\ set\ up\ to\ t) = \exp\left(\alpha IMP_{jkt}^{CH} + f_{kt} + SIC4_j + a_1 \overline{PAT}_{ijk0} + a_2 \overline{IMP}_{jk0}^{CH}\right) \quad (3)$$

Equation (3) can be estimated by either Negative Binomial or Poisson, as in the nonlinear panel models with sequentially exogenous regressors of Blundell et al (1999, 2002).

The estimator used in column (2) of Table 3 does not use initial Chinese imports (i.e. it sets $a_2 = 0$ in equation (3)) so it may not a sufficient approximation for the fixed effect to remove the bias on α in equation (1).[1] We measure initial Chinese imports ($\overline{IMP}_{jk0}^{CH}$) as the average IMP_{jkt}^{CH} across all years from 1990 (our first year of comprehensive imports data) to the year in which a firm enters the sample. For example, the first year of our estimating sample is 1996, so $\overline{IMP}_{jk0}^{CH}$ is the

[1] As Blundell et al. (1999, 2002), note, the bias on the estimate of α converges to zero as the length of pre-sample innovation process becomes long. However, one of the conditions for this asymptotic result is that the fixed effect in the PAT_{ijkt} is proportional to the fixed effect in the IMP_{jkt}^{CH} process. If this is not the case, then it may also be necessary to condition on $\overline{IMP}_{jk0}^{CH}$.

average of IMP_{jkt}^{CH} between 1990 and 1996. For a firm who entered in 1997, $\overline{IMP}_{jk0}^{CH}$ is the 1990-1997 average, and so on. Column (3) of Table 3 includes this measure of $\overline{IMP}_{jk0}^{CH}$ in the specification of the previous column. The coefficient is negative and statistically significant. It is clear that once we control for this initial value of Chinese imports, there is a positive and significant association of innovation with Chinese imports. The significance level (10% level) is lower than in column (1), but the magnitude of the coefficient is larger (1.1 vs. 0.4).

A concern might be that some of the variation in initial Chinese imports is across firms within an industry-country cell. There are two reasons for this. First, we have such variation for the current Chinese import share because some firms operate across multiple industries. For these multi-product firms, we use a weighted average of Chinese import share across all the four digit sectors in which they operate (see BDVR Appendix A2). As an alternative definition, we can allocate a firm solely to its main industry, which is what we do for the rest of the Table 2 for the both the current Chinese import term and its initial condition (labelled "SINGLE" vs. the baseline "MULTI"). Second, Table 3 defines the initial condition as the average Chinese import share between 1990 and the first year we observe the firm in our sample. For firms alive in 1996, it is the 1990-1996 average. However, as noted above, for later entrants we use a longer average as in equation (2): 1997 entrants have the 1990-1997 average, 1998 entrants have the 1990-1998 average, 1990-1998 average for 1999 entrants and the 1990-2000 average for 2000 entrants. We experiment with turning this source of variation off, so that initial Chinese imports is defined solely on the 1990-1996 period for all firms. We label this "FIXED" as opposed to the baseline "COHORT".

We implement these two changes in column (4) of Table 3 that reproduces column (3) but uses a single-industry per firm and define Chinese import initial condition fixed solely in 1990-1996. The coefficient on Chinese imports is 1.087 and significant at the 5% level, near identical to the previous column. Note that the initial imports variable is not statistically significant. This is likely because the initial condition is no longer "initial" for firms who enter after 1996. Since it is the same (the 1990-1996 average) for all firms, it will be a worse control for later entrants.[2] To examine this, column (5) uses the same initial condition approach ("COHORT") as in our baseline

[2] For example, for firms who entered in 2000 (the last entering cohort), the initial condition is 1990-2000 in columns (1) and (2), but 1990-1996 in columns (5) and (6).

models, but continues to allocate firms to a single industry (as in column (3)). As expected, the point estimate on Chinese imports is slightly larger, and the initial conditions are now more precisely estimated. Finally, since equation (3) should also hold if we estimate a Poisson model instead of Negative Binomial model, we repeat the new specifications of Table 3 for the Poisson model, which shows similar qualitative results.[3]

Conclusions

In BDVR, we argued that Chinese import competition played a positive role in upgrading technology in European firms between 2000 and 2007. This conclusion was based on many underlying empirical results showing Chinese competition both reallocated activity to higher tech firms (e.g. reducing employment by more for low-tech firms than for high tech firms) and increased technological change within firms when we examine patents, productivity and IT. CM argue the within firm impact of Chinese imports on patents is sensitive to specification choice. It is true that changing controls can lead to different results on signs and significance, and a useful aspect of our engagement with CM has been to probe the results further in several dimensions, especially of the count data models. Nonetheless, in our view the overall findings from our original paper remain robust when we apply the appropriate corrections.

[3] These are in online Appendix Table A1. Although the Negative Binomial relaxes the distributional assumptions on the error term compared to the simpler Poisson model (it allows for over-dispersion), the fact we cluster the standard errors at the industry-country level means that there is no generality gained by moving from Poisson to NEGBIN (both have the same log-link first moment of equation (3)).

Table 1: Difference in Patent Growth between "Treated" and "Control" Industries

	Sample	Average Patents per firm (base period)	Average Patents per firm (end period)	% change
China competing firms	Full: (1999,2000)–(2004, 2005)	0.681	0.152	-78%
Other firms	Full: (1999,2000)–(2004, 2005)	2.105	0.267	-87%
China competing firms	First half: 1999-2004	0.646	0.261	-60%
Other firms	First half: 1999-2004	2.258	0.416	-82%
China competing firms	Second half: 2000-2005	0.714	0.042	-94%
Other firms	Second half: 2000-2005	1.963	0.112	-94%

Notes: These are on firms in the textiles and clothing industries. Full sample indicates the growth measure uses average patents calculated by pooling 1999 and 2000 for the base period and 2004 and 2005 for the end period. These span China joining the WTO in December 2001.

Table 2: Alternative Functional Forms of the dependent variable in Long-Differences

	(1)	(2)	(3)
Dependent variable format	**Baseline**	**DHS**	**IHS**
Panel A: Full Sample, OLS			
Change in Chinese Imports	0.321**	0.457**	0.400**
	(0.102)	(0.164)	(0.129)
Observations	30,277	30,277	30,277
Panel B: Quota Samples, IV			
Change in Chinese Imports	1.864*	2.665*	2.273*
	(1.008)	(1.480)	(1.270)
Observations	3,443	3,443	3,443

Notes: **indicates significance at the 5% level and * at the 10% level. Column (1) reproduces the results in BDVR (2016) where the dependent variable is $\Delta log(PAT+1)$, where Δ is a 5-year long difference and PAT is the number of patents. Panel A is the full sample (column (1) of Table 1 Panel A in BDVR); Panel B is the OLS results in the quota sample (column (1) of Table 2 in BDVR) and panel C is the IV results (column (3) of Table 2 in BDVR). The F-statistic on the first stage is 24 for Panel B in all columns. Column (2) uses the Davis, Haltiwanger and Shuh (1996) transformation of the dependent variable: $DHS = (PAT_t - PAT_{t-5})/(PAT_t + PAT_{t-5})$ or $DHS = 0$ if $PAT_t = 0$ and $PAT_{t-5} = 0$. Column (3) uses the (long difference of) the Inverse Hyperbolic Sine: $IHS = \Delta log(PAT+ \sqrt{(PAT^2 + 1)})$.

Table 3: Negative Binomial Count Data models with controls for initial Chinese Imports

Dependent Variable:	(1) **PAT+1**	(2) **PAT**	(3) **PAT**	(4) **PAT**	(5) **PAT**
Estimation Method	**NEGBIN**	**NEGBIN**	**NEGBIN**	**NEGBIN**	**NEGBIN**
Current Chinese Imports	0.398**	0.116	1.089*	1.087**	1.350***
	(0.168)	(0.490)	(0.575)	(0.483)	(0.508)
Initial Chinese Imports			-5.371***	-0.403	-1.725*
			(1.401)	(0.934)	(0.953)
Controls	Country and year dummies	Country by year dummies	Country by year dummies	Country by year dummies	Country by year dummies
Definition of Chinese Imports for Current and initial level	MULTI: Average across a firm's industries	MULTI: Average across a firm's industries	MULTI: Average across a firm's industries	SINGLE: Allocated to a single industry	SINGLE: Allocated to a single industry
Timing of initial Chinese imports control	Variable not included	Variable not included	COHORT: Average from 1990 to when firm enters sample	FIXED: Average from 1990-96 for all firms	COHORT: Average from 1990 to when firm enters sample
Observations	74,038	74,038	74,038	74,038	74,038

Notes: ***indicates significance at the 1% level, **5% level and * at the 10% level. PAT is a firm's count of patents. Column (1) is identical to BDVR Table 7 column (1). Column (2) is identical to CM Table 2 column (3). The sample covers the years 1996-2005. All columns include four-digit SIC industry dummies and the two initial condition controls for patents and estimated by Negative Binomial models. Standard errors clustered by industry-country pair. "Current Chinese imports" is the share of Chinese imports in total imports in the industry-country-year cell. In the columns labelled "SINGLE: Allocated to a single industry" we allocate current and initial Chinese imports to the main four digit SIC industry that a firm operates in. "MULTI: Average across a firm's industries" takes into account that some firms operate across multiple industry and uses a weighted average across these industries (as in the original BDVR paper). "Initial Chinese Imports" is the value of the *initial* Chinese import share with the exact timing of this differing by columns. Columns labelled "FIXED: Average from 1990-96 for all firms" uses the average between 1990 and 1996 (so is identical for all firms in a country-industry cell). Columns labelled "COHORT: Average from 1990 to when firm enters sample" uses the 1990-1996 average for firms who were alive in 1996 (i.e. entered the sample in 1996 or earlier); the 1990-1997 average for 1997 entrants, etc.

References

Bloom, Nicholas, Draca, Mirko and Van Reenen, John (2016), "Trade induced technical change? The Impact of Chinese Imports on Innovation, IT and Productivity, *Review of Economic Studies 83(1), 87-117.*

Blundell, Richard, Rachel Griffith and Frank Windmeijer (2002) "Individual effects and Dynamics in Count Data Models" *Journal of Econometrics*, 108(1), 113-131

Blundell, Richard, Rachel Griffith and John Van Reenen (1999). "Market share, Market value and Innovation: Evidence from British Manufacturing Firms" *Review of Economic Studies* 66(3), 228, 529-554

Burbidge, John, Lonnie Magee and A. Leslie Robb (1988) "Alternative Transformations to Handle Extreme Values of the Dependent Variable" *Journal of the American Statistical Association*, 83(401), 123-127

Campbell, Douglas and Karsten Mau (2020) "On Trade Induced Technical Change: The impact of Chinese Imports on Innovation, IT and Productivity", mimeo

Davis, Steven, John Haltiwanger and Scott Schuh (1996) "Small business and job creation: Dissecting the myth and reassessing the facts" *Small Business Economics*, 8(4), 297–31

Appendix Table 1: Patent Count Data models with controls for initial Chinese Imports: Robustness

	(1)	(2)	(3)
Estimation Method:	**POISSON**	**POISSON**	**POISSON**
Current Chinese Imports	2.142** (0.930)	1.937*** (0.634)	2.244*** (0.650)
Initial Chinese Imports	-10.040*** (3.391)	-4.399** (2.007)	-5.632*** (1.974)
Controls	Country by year dummies	Country by year dummies	Country by year dummies
Definition of Chinese Imports for Current and initial level	MULTI: Average across a firm's industries	SINGLE: Allocated to a single industry	SINGLE: Allocated to a single industry
Timing of initial Chinese imports control	COHORT: Average from 1990 to when firm enters sample	FIXED: Average from 1990-96 for all firms	COHORT: Average from 1990 to when firm enters sample
Observations	74,038	74,038	74,038

Notes: ***indicates significance at the 1% level, **5% level and * at the 10% level. PAT is a firm's count of patents. The specification are identical to Table 3 columns (3), (4) and (5), except they are estimated by Poisson instead of a Negative Binomial model. The sample covers the years 1996-2005. All columns include four-digit SIC industry dummies and the two initial condition controls for patents. Standard errors are clustered by industry-country pair. "Current Chinese imports" is the share of Chinese imports in total imports in the industry-country-year cell. In the columns labelled "SINGLE: Allocated to a single industry" we allocate current and initial Chinese imports to the main four digit SIC industry that a firm operates in. "MULTI: Average across a firm's industries" takes into account that some firms operate across multiple industry and uses a weighted average across these industries (as in the original BDVR paper). "Initial Chinese Imports" is the value of the *initial* Chinese import share with the exact timing of this differing by columns. Columns labelled "FIXED: Average from 1990-96 for all firms" uses the average between 1990 and 1996 (so is identical for all firms in a country-industry cell). Columns labelled "COHORT: Average from 1990 to when firm enters sample" uses the 1990-1996 average for firms who were alive in 1996 (i.e. entered the sample in 1996 or earlier); the 1990-1997 average for 1997 entrants; 1990-1998 average for 1998 entrants; the 1990-1998 average for 1999 entrants and the 1990-2000 average for 2000 entrants.