# Supplementary Appendix (Not For Publication) *

October 21, 2017

In Section A of this Internet Appendix, we develop a model of competition with vertical differentiation that allows us to contrast the effect of R&D on firm performance through higher vertical differentiation vs. higher productivity.

Section B includes additional material on the theoretical framework outlined in Section I of the paper:

- analysis of OLS and IV estimators when R&D and trade flows are endogenous (Section B.1);
- extension of the model to the case where $T$ is correlated with $(\theta, \alpha, \beta, \gamma)$ (Section B.2);
- extension of the model to the case where $\mu$ and $\mu^*$ are correlated (Section B.3).

Section C reports additional empirical analysis and robustness tests:

- results when import penetration is scaled with ten-year-lagged employment (Section C.1);
- test of endogeneity of state R&D tax credit policy to economic conditions (Section C.2);
- results when the measure of firms' exposure to state tax credits is based on the location of their headquarters (Section C.3);

*Hombert, Johan and Adrien Matray, Internet Appendix to "Can Innovation Help U.S. Manufacturing Firms Escape Import Competition from China?" Journal of Finance [DOI STRING]. Please note: Wiley-Blackwell is not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the authors of the article.

- results when the measure firms' exposure to import penetration based on Compustat Business Segment data (Section C.4);
- results when excluding California (Section C.5);
- tests of non-monotonicity of the effect import competition on returns to R&D (Section C.6);
- analysis of product differentiation using our second proxy of differentiation relative to Chinese competitors (Section C.7).

# A IO Model

There are two ways by which innovation can translate into higher product market performance. The one traditionally considered in the innovation and growth literature is that innovation leads to an increase in productivity, as Grossman and Helpman (1991) and Aghion and Howitt (1992). Second, innovation may improve product market performance through vertical differentiation, as analyzed by Sutton (1991).

The distinction between the two types of innovation is important because they lead to different empirical implications. In this appendix, we work out a textbook model of competition with vertical differentiation showing that, while both types of innovation lead to an unconditional increase in firm profit, they have opposite effects *conditional* on the intensity of competition. Importantly, the distinction between productivity improvement and vertical differentiation has relevant empirical content. Hoberg and Phillips (2015) develop a methodology to measure the similarity between firms' products based on the textual analysis of product descriptions in firms' 10-K, which can be used to identify industries in which products are more homogeneous (and innovation is more about improving productivity) and industries in which products are more differentiated (and innovation is more about increasing vertical differentiation).

**Setup** [1] There is a mass one continuum of consumers with heterogeneous valuation $\xi$ for quality. $\xi$ is uniformly distributed on $[0, 1]$. Each consumer consumes one or zero units of a good. A consumer with valuation $\xi$ for quality derives utility $\xi q - p$ from purchasing a good of quality $q$ at price $p$, or $u_0$ if he does not purchase the good. We assume that $u_0$ is low enough such that all consumers purchase the good in equilibrium.

There is one domestic firm and a competitive fringe of foreign firms. The domestic firm and foreign firms have marginal cost $c$ and $c^*$, respectively, and they produce goods of quality $q$ and $q^*$, respectively. We assume that the domestic firm is initially more efficient than foreign firms on both the productivity dimension and the quality dimension: $c < c^*$ and $q > q^*$. To focus on interesting cases, we also assume that the cost differential between domestic and foreign firms is not too large to ensure that the demand addressed to both types of firms is nonzero in equilibrium: $c^* - c < q - q^*$.

[1] The setup follows closely the presentation in Tirole (1988, Chapter 2).

Since foreign firms are competitive and produce homogeneous products, they charge a price $p^* = c^*$. The domestic firm faces the demand function:

$$D(p) = \begin{cases} 1 & \text{if } p \leq c^*, \\ 1 - \frac{p-c^*}{q-q^*} & \text{if } p \in [c^*, c^* + q - q^*], \\ 0 & \text{if } p \geq c^* + q - q^*. \end{cases}$$

Its profit $\pi(p) = (p - c)D(p)$ is maximized for a price $p \in [c^*, c^* + q - q^*]$ such that the first order condition $\pi'(p) = 0$ holds. It implies $p = \frac{1}{2}\big((q - q^*) + (c + c^*)\big)$ and equilibrium profit for the domestic firm is equal to:

$$\Pi(c, q, c^*, q^*) = \frac{\big((q - q^*) + (c^* - c)\big)^2}{4\big(q - q^*\big)}. \tag{A.1}$$

$\Pi(c, q, c^*, q^*)$ is decreasing in $c$ and increasing in $q$. Thus, both productivity-enhancing innovation and vertical differentiation-enhancing innovation lead to higher performance for the domestic firm.

**Increase in import competition** We model an increase in import competition as a reduction in foreign firms' marginal cost $c^*$. The impact on the performance of the domestic firm is given by:

$$-\frac{\partial \Pi}{\partial c^*} = -\frac{(q - q^*) + (c^* - c)}{2(q - q^*)} < 0. \tag{A.2}$$

Thus, import competition weighs on the domestic firm's profits. More important to us is whether the negative effect of foreign competition is stronger or weaker when the domestic firm has done more productivity-enhancing innovation (i.e., it has lower $c$) and when it has done more vertical differentiation-enhancing innovation (i.e., it has higher $q$).

Consider first the case of productivity-enhancing innovation. The adverse effect of import competition given by (A.2) is stronger (more negative) when $c$ is lower. It reflects the Schumpeterian effect by which the benefit of higher productivity is eroded by competition. In the language of the model presented in Section I, $\delta$ is negative for productivity-enhancing innovation.

In the case of vertical differentiation-enhancing innovation, the adverse effect of import competition given by (A.2) is lower (less negative) when $q$ is higher. It reflects the effect by which vertical differentiation allows firms to escape import competition. Thus, $\delta$ is positive for vertical differentiation-enhancing innovation.

These results have the following implications:

**Implication A.1.** *An increase in competition has a less adverse effect on firms with more differentiated products.*

We test Implication A.1 in Section IV.A of the paper.

**Implication A.2.** *If R&D allows firms to choose between increasing productivity or increasing differentiation, then an increase in competition shifts the optimal choice towards increasing differentiation.*

We test Implication A.2 in Section IV.B of the paper. The last implication relies on the idea that the extent to which innovation allows firms to increase productivity or to increase differentiation varies across industries.

**Implication A.3.** $\delta$ *is higher (more positive) in industries in which innovation is differentiation-enhancing relative to industries in which innovation is productivity-enhancing.*

We test Implication A.3 in Section IV.C of the paper, where we proxy for the importance of vertical differentiation as the average distance between firms' products in the industry.

# B Additional Material on the Theoretical Framework

This appendix presents additional material on the theoretical framework outlined in Section I of the paper.

## B.1 OLS and IV Estimators

In this section, we study the biases arising in the OLS estimator of equation (5) because of the endogeneity of R&D and trade flows. To isolate each source of bias and study to correct each of them, we proceed in two steps. In Section B.1.1, we analyze the biases coming from endogenous R&D (abstracting from endogenous trade shocks) and show how to correct them by instrumenting for R&D. In Section B.1.2, we analyze the bias arising when trade flows respond endogenously to domestic shocks and show how to correct it by instrumenting for imports. Finally, to facilitate the economic interpretation, we compute first-order Taylor approximations for small second-order moments of $T$, which enables us to obtain closed-form expressions for all the estimators of $\delta$ we analyze. All the proofs are relegated in Section B.1.3.

### B.1.1 Endogenous R&D

To focus on the issue of endogenous R&D, we first consider the case where the econometrician can observe a measure of import competition ($T$) that is not correlated with $(\theta, \alpha, \beta, \gamma)$.[2] The following proposition analyzes the OLS estimator of equation (5) in the paper when endogenous R&D ($R$) is used as a regressor.

**Proposition B.1.** *The expected estimator of $\delta$ in (5) when endogenous R&D is used as a regressor is:*

$$E\Big[\hat{\delta}^{R\&D-ols}\Big] = \delta\Big(1 + \frac{Cov(\gamma, \gamma-\theta)}{V(\gamma-\theta)}\Big) + \frac{Cov(\beta, \gamma-\theta)}{V(\gamma-\theta)}\rho. \tag{B.1}$$

Proposition B.1 shows that two biases arise in the OLS estimator. To highlight the economic intuition, we discuss here the case where $\theta$ is constant, in which case the

[2]The case where $T$ is correlated with $(\theta, \alpha, \beta, \gamma)$ is analyzed in Section B.2.

expected OLS estimator is equal to $E\left[\hat{\delta}^{R\&D-ols}\right] = 2\delta + \dfrac{Cov(\beta,\gamma)}{V(\gamma)}\rho$.[3]

The first bias is that $\delta$ is estimated with an inflated factor of two. It arises because of unobserved heterogeneity in $\gamma$. The intuition is the following. Firms with a high benefit of innovation (high $\gamma$) do more R&D and, as a result, have higher performance. These firms also have higher performance at any level of R&D because they have higher returns to R&D. Thus, OLS estimates the sensitivity of performance to R&D with an upward bias. Because the bias is driven by unobserved heterogeneity in $\gamma$, the magnitude of the bias is proportional to the marginal effect of $\gamma$ on performance, i.e., the bias is proportional to R&D. Now, $\delta$ measures how the performance-R&D sensitivity depends on import competition. If $\delta > 0$, industries exposed to higher competition have higher R&D and thus the estimated performance-R&D sensitivity is estimated with a larger upward bias in these industries relative to industries less exposed to import competition. Thus, $\delta$ is estimated with an upward bias. Conversely, if $\delta < 0$, industries exposed to higher competition have lower R&D and thus the estimated performance-R&D sensitivity is estimated with a smaller upward bias in these industries. Thus, $\delta$ is estimated with a downward bias. In both cases, the OLS estimator of $\delta$ is biased away from zero.

The second bias arises when the resilience to trade shocks ($\beta$) is correlated with the benefit of innovation ($\gamma$). For instance, if firms that are better managed are more resilient to trade shocks and also do more R&D, then there will be a spurious positive correlation between R&D and resilience to trade shocks that does not reflect the causal effect of R&D.

The next proposition shows that these biases can be corrected by instrumenting for R&D. Suppose we have at our disposal a variable $z$ that shifts the cost of R&D ($\theta$) and is orthogonal to other exogenous variables. In our empirical analysis, $z$ is an R&D tax credit. Proposition B.2 analyzes the IV estimator of equation (5) when we instrument $(T, R, RT)$ using $(T, z, zT)$.

**Proposition B.2.** *The expected estimator of $\delta$ in (5) when R&D is instrumented by an exogenous cost shifter is:*

$$E\left[\hat{\delta}^{R\&D-iv}\right] = \delta. \tag{B.2}$$

[3] Similar mechanisms are at work with heterogeneous $\theta$.

Thus, the biases in the OLS estimator stemming from the endogeneity of the R&D decision are eliminated when R&D is instrumented by an exogenous cost shifter.

### B.1.2 Endogenous trade flows

In the previous section, the econometrician was assumed to observe directly productivity shocks in the foreign country. We now analyze the case where the econometrician can only observe trade flows from the foreign country to the domestic country and try to infer from them the underlying productivity shocks in the foreign country.

In any theory of international trade, trade flows depend both on productivity in the exporting country and on productivity in the domestic country. The latter can threaten the identification of $\delta$, because innovation shocks in the domestic country can generate a spurious correlation between imports and the performance of innovative firms. Intuitively, positive innovation shocks in the domestic country raises aggregate domestic productivity and thus lowers imports, and they also increase the productivity of innovative firms.

More formally, the econometrician observes imports at the industry level. Imports depend on the gap between foreign productivity and domestic productivity. In our simple framework, we proxy for domestic productivity $A_j$ in industry $j$ as the average performance that firms in this industry would experience in the absence of international trade (i.e., setting $T_j = 0$ in equation (3)). Thus, $A_j = \frac{1}{|j|}\sum_{i\in j}\left(\alpha_i + \gamma_i I_i\right)$ and imports in industry $j$ are equal to:

$$Imports_j = T_j - \lambda A_j, \tag{B.3}$$

where $\lambda > 0$ parameterizes the sensitivity of imports to domestic productivity.

We assume that shocks to the innovation outcome do not perfectly average out at the industry level: $\frac{1}{|j|}\sum_{i\in j} I_i \neq \frac{1}{|j|}\sum_{i\in j} E[I_i|R_i]$. This assumption captures the notion that industries are granular, breaking the Law Of Large Numbers (LOLN) (Gabaix (2011)). If the LOLN held, aggregate innovation at the industry level would be a deterministic function of aggregate R&D in the industry. From an econometric perspective, it would mean that R&D is a perfect measure of innovation at the industry level. Instead, we assume that the LOLN does not apply, which implies that innovation can vary across industries even for the same level of R&D. Formally, we assume there are industry-level innovation shocks, $\mu_j$, that make firm-level innovation correlated with industry-level innovation conditional on R&D:[4]

$$\text{for } i \in j, \quad P[I_i = 1 | R_i, \mu_j] = \mu_j R_i. \tag{B.4}$$

Finally, to focus on the endogeneity of trade flows in relation with innovation shocks, we make the simplifying assumption that the distribution of $(\alpha, \beta, \gamma, \theta)$ across firms is the same in all industries.

The following proposition calculates the estimator of $\delta$ equation (5) when R&D is instrumented by an exogenous cost shifter (such that endogeneity of R&D is no longer a problem) and imports are used as a proxy for foreign productivity shocks ($T$).

**Proposition B.3.** *The expected estimator of $\delta$ in (5) when R&D is instrumented by an exogenous cost shifter and foreign productivity is measured using domestic imports is:*

$$E\left[\hat{\delta}^{import-ols}\right] = \delta - \kappa\,\lambda\,V(\mu), \tag{B.5}$$

*where $\kappa > 0$ is reported in equation (B.9).*

A first observation from Proposition B.3 is that, when imports do not depend on domestic productivity ($\lambda = 0$), we are back to the case where the estimator is unbiased. However, when imports depend on domestic productivity ($\lambda > 0$), the OLS estimator is downward biased. When domestic firms successfully innovate (high realized $\mu_j$), the realized returns to R&D are large and, at the same time, imports are low because domestic productivity is high relative to foreign productivity. This mechanism generates a spurious negative relation between realized returns to R&D and realized imports, creating a downward bias in the estimate of $\delta$.

The next proposition shows that this bias can be eliminated by extracting foreign productivity shocks from foreign imports to a third country that has a similar economic structure as the domestic country. Consider a third country (or group of third countries) described by equations mirroring the ones for the domestic country. Namely, in the

[4]Since $\mu_j$ is interpreted as a (non-zero) average of idiosyncratic shocks that are assumed to be independent from other exogenous variables, we assume that $\mu_j$ is also independent from other exogenous variables. It then follows from $P[I_i = 1|R_i] = R_i$ and Bayes law that $E[\mu_j] = 1$.

third country, firm performance is given by equation (3), firm innovation is given by equation (B.4) where $\mu_j$ is replaced by $\mu_j^*$ representing the part of the idiosyncratic shocks that fails to average out across firms belonging to industry $j$ in the third country, and imports by the third country ($Imports_j^*$) are given by equation (B.3) where $A_j$ is replaced by $A_j^*$ defined as before as average firm performance in the third country's industry $j$ that would prevail in the absence of international trade.

Proposition B.4 analyzes the estimator of equation (5) when $(Imports, R, R.Imports)$ is instrumented using $(Imports^*, z, z.Imports^*)$ and innovation shocks in the domestic country and the third country are not correlated.[5]

**Proposition B.4.** *Assume $\mu$ and $\mu^*$ are not correlated. The expected estimator of $\delta$ in (5) when R&D is instrumented by an exogenous cost shifter and domestic imports are instrumented by the third country's imports is:*

$$E\left[\hat{\delta}^{import-iv}\right] = \delta. \tag{B.6}$$

Proposition B.4 shows that the downward bias stemming from the endogeneity of trade flows to innovation shocks in the destination country is eliminated when we use imports to a third country to estimate the returns to innovation in the domestic country. Since the bias was always downwards in Proposition B.3, it implies that using imports by the third country as a regressor leads to a higher (and unbiased) estimate of $\delta$ than when using imports by the domestic country.

### B.1.3 Proofs

**Notations** In all the proofs, for any random variables $x$ and $y$, we denote $\bar{x} \equiv E[x]$, $\sigma_x^2 \equiv Var(x)$, and $\sigma_{xy} \equiv Cov(x, y)$. $x = o(\varepsilon)$ means that $x/\varepsilon$ goes to zero when $\varepsilon$ goes to zero (i.e., $x$ is of an order of magnitude smaller than $\varepsilon$). $y = O(\varepsilon)$ to mean that $y/\varepsilon$ is bounded above when $\varepsilon$ goes to zero (i.e., $y$ is at most of the same order of magnitude as $\varepsilon$).

[5] The case where $\mu$ and $\mu^*$ are correlated is analyzed in Section B.3.

**Proof of Proposition B.1** The OLS estimator of $(\beta, \gamma, \delta)'$ in equation (5) is equal to $Cov(\mathbf{X}', \mathbf{X})^{-1} Cov(\mathbf{X}', \pi)$, where $\mathbf{X} \equiv (R, T, RT)'$, $R$ is given by (4), and $\pi$ is given by (3). The random variables in $\mathbf{X}$ and $\pi$ are $\zeta \equiv (\alpha, \beta, \gamma, \theta, T)$ and $I$. It follows from the law of total covariance that $Cov\big(\mathbf{X}', \pi\big) = Cov\big(E[\mathbf{X}'|\zeta], E[\pi|\zeta]\big) + E\big[Cov(\mathbf{X}', \pi|\zeta)\big]$. Since $\mathbf{X}|\zeta$ is a constant, the second term is equal to zero and, in the first term, $\mathbf{X}'|\zeta = \mathbf{X}'$. Thus, denoting the expected performance conditional on $\zeta$ by $Y \equiv E[\pi|\zeta]$, the OLS estimator is equal to $Cov(\mathbf{X}', \mathbf{X})^{-1} Cov(\mathbf{X}', Y)$.

In order to obtain closed-form expressions for the estimators, we consider small second-order moments of $T$. First, we can normalize $\bar{T} = 0$ without affecting the value of $\delta$. Then, we assume that $\sigma_T^2$ is small and that $E[T^3] = o(\sigma_T^2)$, $E[T^4] = o(\sigma_T^2)$, and for $x \in \{\alpha, \beta, \gamma, \theta\}$, $x = \bar{x} + \frac{\sigma_{xT}}{\sigma_T^2} T + \varepsilon_x$ where $\sigma_{xT} = O(\sigma_T^2)$ and $\varepsilon_x$ is independent from $T$. A useful formula implied by these assumptions is that, for $(x, y, z) \in \{\alpha, \beta, \gamma, \theta, T\}^3$, $Cov(xy, zT) = \big(\bar{x}\bar{z} + \sigma_{xz}\big)\sigma_{yT} + \big(\bar{y}\bar{z} + \sigma_{yz}\big)\sigma_{xT} + o(\sigma_T^2)$.[6]

Using

$$
\begin{aligned}
R &= \frac{1}{\rho}\big(\gamma + \delta T - (1+\theta)\big) \\
Y &= \alpha + \beta T + \frac{1}{\rho}\big(\gamma + \delta T - (1+\theta)\big)\big(\gamma + \delta T\big),
\end{aligned}
$$

[6]Proof: $Cov(xy, zT) = E[xyzT] - E[xy]E[zT] = E\big[(\bar{x} + \varepsilon_x)(\bar{y} + \varepsilon_y)\frac{\sigma_{zT}}{\sigma_T^2} T^2 + (\bar{x} + \varepsilon_x)(\bar{z} + \varepsilon_z)\frac{\sigma_{yT}}{\sigma_T^2} T^2 + (\bar{y} + \varepsilon_y)(\bar{z} + \varepsilon_z)\frac{\sigma_{xT}}{\sigma_T^2} T^2\big] + o(\sigma_T^2) - E[xy]\sigma_{xT} = \big(\bar{x}\bar{z} + \sigma_{xz}\big)\sigma_{yT} + \big(\bar{y}\bar{z} + \sigma_{yz}\big)\sigma_{xT} + o(\sigma_T^2)$.

we calculate:

$$
\begin{aligned}
V(R) &= \frac{1}{\rho^2}\Big(\sigma_\gamma^2 + \sigma_\theta^2 - 2\sigma_{\gamma\theta}\Big) + o(1),\\
Cov(R,T) &= \frac{1}{\rho}\delta\sigma_T^2,\\
Cov(R,RT) &= \frac{\bar{\gamma} - 1 - \bar{\theta}}{\rho^2}\delta\sigma_T^2 + o(\sigma_T^2),\\
V(T) &= \sigma_T^2,\\
Cov(T,RT) &= \frac{\bar{\gamma} - 1 - \bar{\theta}}{\rho}\sigma_T^2 + o(\sigma_T^2),\\
V(RT) &= \frac{1}{\rho^2}\Big((\bar{\gamma} - 1 - \bar{\theta})^2 + \sigma_\gamma^2 + \sigma_\theta^2 - 2\sigma_{\gamma\theta}\Big)\sigma_T^2 + o(\sigma_T^2),\\
Cov(Y,R) &= \frac{1}{\rho}\Big(\sigma_{\alpha\gamma} - \sigma_{\alpha\theta}\Big) + \frac{1}{\rho^2}\Big((2\bar{\gamma} - 1 - \bar{\theta})\sigma_\gamma^2 + \bar{\gamma}\sigma_\theta^2 + (1 + \bar{\theta} - 3\bar{\gamma})\sigma_{\gamma\theta}\Big) + o(1),\\
Cov(Y,T) &= \Big(\bar{\beta} + \delta\frac{2\bar{\gamma} - 1 - \bar{\theta}}{\rho}\Big)\sigma_T^2 + o(\sigma_T^2),\\
Cov(Y,RT) &= \Big(\frac{\bar{\beta}(\bar{\gamma} - 1 - \bar{\theta})}{\rho} + \frac{1}{\rho}\big(\sigma_{\beta\gamma} - \sigma_{\beta\theta}\big) + \frac{\delta}{\rho^2}\big((2\bar{\gamma} - 1 - \bar{\theta})(\bar{\gamma} - 1 - \bar{\theta}) + 2\sigma_\gamma^2 + \sigma_\theta^2 - 3\sigma_{\gamma\theta}\big)\Big)\sigma_T^2\\
&\quad + o(\sigma_T^2).
\end{aligned}
$$

We calculate $Cov(\mathbf{X}', \mathbf{X})Cov(\mathbf{X}', Y)$ using Mathematica and obtain the OLS estimator of $\delta$:

$$
E\Big[\hat{\delta}^{R\&D-ols}\Big] = \delta\Big(1 + \frac{Cov(\gamma, \gamma - \theta)}{V(\gamma - \theta)}\Big) + \frac{Cov(\beta, \gamma - \theta)}{V(\gamma - \theta)}\rho + o(1). \tag{B.7}
$$

**Proof of Proposition B.2** The instrument $z$ is such that $\theta = \tilde{\theta} + z$ and $z$ has zero mean and is orthogonal to $(\alpha, \beta, \gamma, \tilde{\theta}, T)$. We denote the vector of instruments by $\mathbf{Z} \equiv (z, T, zT)'$. Applying the law of total covariance as in the proof of Proposition B.1, the IV estimator is equal to $Cov(\mathbf{Z}', \mathbf{X})^{-1}Cov(\mathbf{Z}', \pi) = Cov(\mathbf{Z}', \mathbf{X})^{-1}Cov(\mathbf{Z}', Y)$, where $Y \equiv E[\pi|\zeta]$ denotes the expected performance conditional on $\zeta \equiv (\alpha, \beta, \gamma, \tilde{\theta}, z, T)$.

Using

$$
\begin{aligned}
R &= \frac{1}{\rho}\big(\gamma + \delta T - (1 + \tilde{\theta} + z)\big)\\
Y &= \alpha + \beta T + \frac{1}{\rho}\big(\gamma + \delta T - (1 + \tilde{\theta} + z)\big)\big(\gamma + \delta T\big),
\end{aligned}
$$

we calculate:

$$
\begin{aligned}
Cov(z,R) &= -\frac{1}{\rho}\sigma_z^2,\\
Cov(z,T) &= 0,\\
Cov(z,RT) &= 0,\\
Cov(T,R) &= \frac{\delta}{\rho}\sigma_T^2 + o(\sigma_T^2),\\
Cov(T,T) &= \sigma_T^2,\\
Cov(T,RT) &= \frac{\bar{\gamma}-1-\bar{\theta}}{\rho}\sigma_T^2 + o(\sigma_T^2),\\
Cov(zT,R) &= 0,\\
Cov(zT,T) &= 0,\\
Cov(zT,RT) &= -\frac{1}{\rho}\sigma_z^2\sigma_T^2 + o(\sigma_T^2),\\
Cov(Y,z) &= -\frac{\bar{\gamma}}{\rho}\sigma_z^2 + o(\sigma_T^2),\\
Cov(Y,T) &= \Big(\bar{\beta} + \delta\frac{2\bar{\gamma}-1-\bar{\theta}}{\rho}\Big)\sigma_T^2 + o(\sigma_T^2),\\
Cov(Y,zT) &= -\frac{1}{\rho}\sigma_z^2\delta\sigma_T^2 + o(\sigma_T^2).
\end{aligned}
$$

We calculate $Cov(\mathbf{Z}',\mathbf{X})^{-1}Cov(\mathbf{Z}',Y)$ using Mathematica and obtain the IV estimator of $\delta$:

$$E\Big[\hat{\delta}^{R\&D-iv}\Big] = \delta + o(1). \tag{B.8}$$

**Proof of Proposition B.3** We denote $M \equiv \mathit{Imports}$, the vector of covariates by $\mathbf{X} \equiv (R, M, RM)'$, and the vector of instruments by $\mathbf{Z} \equiv (z, M, zM)'$. Applying the law of total covariance as in the proof of Proposition B.1, the IV estimator is equal to $Cov(\mathbf{Z}',\mathbf{X})^{-1}Cov(\mathbf{Z}',\pi) = Cov(\mathbf{Z}',\mathbf{X})^{-1}Cov(\mathbf{Z}',Y)$, where $Y \equiv E[\pi|\zeta]$ denotes the expected performance conditional on $\zeta \equiv (\alpha, \beta, \gamma, \tilde{\theta}, z, T)$.

Letting $\Omega \equiv (\gamma - 1 - \theta)\gamma$, equation (B.3) implies that imports in industry $j$ are equal to $M_j = T_j - \lambda(\bar{\alpha} + \mu_j \bar{\Omega})$. Using

$$\begin{aligned} R &= \frac{1}{\rho}\big(\gamma + \delta T - (1 + \tilde{\theta} + z)\big) \\ M &= T - \lambda(\bar{\alpha} + \mu_j \bar{\Omega}), \\ Y &= \alpha + \beta T + \mu_j \frac{1}{\rho}\big(\gamma + \delta T - (1 + \tilde{\theta} + z)\big)\big(\gamma + \delta T\big), \end{aligned}$$

we calculate:

$$\begin{aligned} Cov(z, R) &= -\frac{1}{\rho}\sigma_z^2, \\ Cov(z, M) &= 0, \\ Cov(z, RM) &= \frac{\lambda(\bar{\alpha} + \bar{\Omega})}{\rho}\sigma_z^2, \\ Cov(zM, R) &= \frac{\lambda(\bar{\alpha} + \bar{\Omega})}{\rho}\sigma_z^2, \\ Cov(zM, M) &= 0, \\ Cov(zM, RM) &= -\frac{1}{\rho}\Big(\lambda^2\big((\bar{\alpha} + \bar{\Omega})^2 + \bar{\Omega}^2\sigma_\mu^2\big) + \sigma_T^2\Big)\sigma_z^2, \\ Cov(Y, z) &= -\frac{\bar{\gamma}}{\rho}\sigma_z^2, \\ Cov(Y, zM) &= -\frac{1}{\rho}\Big(-\bar{\gamma}\lambda(\bar{\alpha} + \bar{\Omega}) + \delta\sigma_T^2 - \bar{\gamma}\lambda\bar{\Omega}\sigma_\mu^2\Big)\sigma_z^2. \end{aligned}$$

We calculate $Cov(\mathbf{Z}', \mathbf{M})^{-1} Cov(\mathbf{Z}', Y)$ and obtain:

$$E\Big[\hat{\delta}^{import-ols}\Big] = \delta - \frac{\Omega(\bar{\gamma} + \delta\bar{\Omega}\lambda)}{\sigma_T^2 + \bar{\Omega}^2\lambda^2\sigma_\mu^2}\lambda\,\sigma_\mu^2 + o(1). \tag{B.9}$$

**Proof of Proposition B.4** We denote $M^* \equiv Imports^*$, the vector of covariates by $\mathbf{X} \equiv (R, M, RM)'$, and the vector of instruments by $\mathbf{Z} \equiv (z, M^*, zM^*)'$. Equation (B.3) implies that imports by the third country are equal to $M^* = T - \lambda(\bar{\alpha} + \mu_j^* \bar{\Omega})$. Following the same steps as in the proof of Proposition B.3, we calculate

$$\begin{aligned}
Cov(z,R) &= -\frac{1}{\rho}\sigma_z^2,\\
Cov(z,M) &= 0,\\
Cov(z,RM) &= \frac{\lambda(\bar{\alpha}+\bar{\Omega})}{\rho}\sigma_z^2,\\
Cov(zM^*,R) &= \frac{\lambda(\bar{\alpha}+\bar{\Omega})}{\rho}\sigma_z^2,\\
Cov(zM^*,M) &= 0,\\
Cov(zM^*,RM) &= -\frac{1}{\rho}\Big(\lambda^2(\bar{\alpha}+\bar{\Omega})^2+\sigma_T^2\Big)\sigma_z^2,\\
Cov(Y,z) &= -\frac{\bar{\gamma}}{\rho}\sigma_z^2,\\
Cov(Y,zM^*) &= -\frac{1}{\rho}\Big(-\bar{\gamma}\lambda(\bar{\alpha}+\bar{\Omega})+\delta\sigma_T^2\Big)\sigma_z^2.
\end{aligned}$$

We calculate $Cov(\mathbf{Z}',\mathbf{M})^{-1}Cov(\mathbf{Z}',Y)$ and obtain:

$$E\Big[\hat{\delta}^{import-iv}\Big]=\delta+o(1). \tag{B.10}$$

## B.2 Import Competition Shocks Correlated with Domestic Shocks

We analyze the case where productivity shocks in the foreign country ($T$) are correlated with firms' opportunity cost of R&D ($\theta$), baseline level of performance ($\alpha$), resilience to trade shocks ($\beta$), and unconditional return to innovation ($\gamma$). Correlation between $T$ and $(\theta,\alpha,\beta,\gamma)$ raises a different set of issues than the case in which the econometrician can only observe realized trade flows as analyzed in Section B.1.2 of this Internet Appendix. The former generates a spurious correlation between *expected* performance and competition, whereas the latter generates a spurious correlation between *realized* performance and competition.

The following proposition analyzes the IV estimator of equation (5) when $(T,R,RT)$ is instrumented using $(T,z,zT)$, as well as the same estimator when we include industry fixed effects interacted with $z$ as an exogenous regressor.

**Proposition B.5.** *If $T$ is correlated with $(\theta,\alpha,\beta,\gamma)$, then the expected estimator of $\delta$ in*

*(5) when R&D is instrumented by an exogenous cost shifter is:*

$$E\left[\hat{\delta}^{R\&D-iv}\right] = \delta + \frac{Cov(\gamma, T)}{V(T)}. \tag{B.11}$$

*If, in addition, the data is panel and $\gamma$ is an industry fixed characteristic, then the expected estimator when industry fixed effects interacted with $z$ are included is:*

$$E\left[\hat{\delta}^{R\&D-iv,indFE.z}\right] = \delta. \tag{B.12}$$

The first part of Proposition B.5 shows that the estimator is biased when trade shocks ($T$) are correlated with the sensitivity of performance to innovation ($\gamma$). If the correlation is, say, positive, then the estimated return to innovation will be larger in import-exposed industries, leading to a higher estimated $\delta$. However, it will only reflect unobserved heterogeneity in the exogenous returns to innovation and not the causal effect of import competition.

The second part of the proposition shows that, if exogenous return to innovation ($\gamma$) is an fixed industry characteristic, then this bias can be corrected by controlling for industry fixed effects interacted with the instrument for R&D ($z$). Intuitively, industry fixed effects interacted with instrumented R&D control for unobserved heterogeneity in time-invariant industry-specific return to innovation ($\gamma$), and thus correct the bias arising from potential correlation between $\gamma$ and $T$.

To check that the results in the paper are not driven by such unobserved heterogeneity, we re-run our main regressions including industry fixed effects interacted with instrumented R&D. Results reported in Table B.1 show that our main results are robust to this alternative specification.

**Proof of Proposition B.5** It follows from the Frisch-Waugh-Lovell theorem that adding industry fixed effects interacted with $z$ as an exogenous regressor is equivalent to projecting $Y$, $(T, RT)$, and $(T, zT)$ on industry fixed effects interacted with $z$ and computing the IV estimator using the residuals. These residuals are equal to:

$$\begin{aligned}
\widetilde{Y} &= \alpha + \beta T + \frac{1}{\rho}(\gamma - 1 - \tilde{\theta} + \delta T)(\gamma + \delta T) - z(\gamma^{\perp} + \delta T^{\perp}), \\
\widetilde{T} &= T, \\
\widetilde{RT} &= (\gamma - 1 - \tilde{\theta} + \delta T)T - zT^{\perp}, \\
\widetilde{zT} &= zT^{\perp},
\end{aligned}$$

where $\gamma^{\perp}$ and $T^{\perp}$ denote the deviation of $\gamma$ and $T$, respectively, from the industry time-series average. We calculate:

$$\begin{aligned}
Cov(\widetilde{T}, \widetilde{T}) &= \sigma_T^2, \\
Cov(\widetilde{T}, \widetilde{RT}) &= \frac{\bar{\gamma} - 1 - \bar{\theta}}{\rho}\sigma_T^2 + o(\sigma_T^2), \\
Cov(\widetilde{zT}, \widetilde{T}) &= 0, \\
Cov(\widetilde{zT}, \widetilde{RT}) &= -\sigma_z^2 \frac{1}{\rho}\sigma_{T^{\perp}}^2 + o(\sigma_T^2), \\
Cov(\widetilde{Y}, \widetilde{T}) &= \sigma_{\alpha T} + (\bar{\beta} + \delta\frac{2\bar{\gamma} - 1 - \bar{\theta}}{\rho})\sigma_T^2 + \frac{2\bar{\gamma} - 1 - \bar{\theta}}{\rho}\sigma_{\gamma T} - \frac{\bar{\gamma}}{\rho}\sigma_{\theta T} + o(\sigma_T^2), \\
Cov(\widetilde{Y}, \widetilde{zT}) &= -\sigma_z^2 \frac{\delta}{\rho}\sigma_{T^{\perp}}^2 + o(\sigma_T^2).
\end{aligned}$$

We calculate the IV estimator $Cov((\widetilde{T}, \widetilde{zT})', (\widetilde{T}, \widetilde{RT}))^{-1} Cov((\widetilde{T}, \widetilde{zT})', Y)$ using Mathematica and obtain:

$$E\left[\hat{\delta}^{R\&D-iv,indFE.z}\right] = \delta + o(1).$$

Table B.1: Controlling for Industry Fixed Effects Interacted with Instrumented R&D

| | Sales growth (1) | ROA (2) | Capital expenditures (3) | Employment growth (4) |
|---|---|---|---|---|
| Import penetration × R&D stock | 0.99** | 1.19** | 2.89*** | 0.72* |
| | (0.49) | (0.60) | (0.66) | (0.42) |
| Assets | 0.04*** | 0.02* | 0.06*** | 0.08*** |
| | (0.01) | (0.01) | (0.02) | (0.01) |
| Age | -0.24*** | 0.15*** | -0.45*** | -0.26*** |
| | (0.02) | (0.02) | (0.03) | (0.01) |
| R&D stock | 0.21 | -0.12 | 0.37 | 0.15 |
| | (0.42) | (0.10) | (0.40) | (0.26) |
| Import penetration × Age | -0.70 | 0.18 | -1.20 | -0.90** |
| | (0.48) | (0.41) | (0.73) | (0.46) |
| Firm FE | Yes | Yes | Yes | Yes |
| Industry-Year FE | Yes | Yes | Yes | Yes |
| Industry FE × R&D | Yes | Yes | Yes | Yes |
| Observations | 23,321 | 23,959 | 23,748 | 22,576 |
| R2 | .34 | .73 | .41 | .35 |

The sample is US manufacturing firms over 1991–2007 from Compustat. We estimate the same regression as in column (4) of Tables VI to XII and additionally include industry fixed effects instrumented with the instrumented R&D capital stock. Standard errors are bootstrapped within industry-year clusters and reported in parentheses. *, **, and *** mean statistically different from zero at 10, 5, and 1% levels of significance.

## B.3 Correlated Innovation Shocks in the Third Country and in the Domestic Country

We analyze the IV estimator of Proposition B.4 (imports in the domestic country instrumented using imports in the third party) in the case where innovation shocks in the domestic country and the third country are correlated: $Cor(\mu, \mu^*) \neq 0$.

**Proposition B.6.** *Assume $\mu$ and $\mu^*$ are correlated. The expected estimator of $\delta$ in (5) when R&D is instrumented by an exogenous cost shifter and domestic imports are instrumented by the third country's imports is:*

$$E\left[\hat{\delta}^{import-iv}\right] = \delta - \kappa\big(Cor(\mu, \mu^*)\big)\,\lambda\,V(\mu), \tag{B.13}$$

*where $\kappa(.)$ is increasing in $Cor(\mu, \mu^*)$, $\kappa(0) = 0$, and $\kappa(1) = \kappa$.*

Proposition B.6 shows that the IV still corrects part of the bias in the OLS estimate as long as innovation shocks in the domestic country and in the third country are not perfectly correlated. The bias is completely eliminated when $Cor(\mu, \mu^*) = 0$, as already shown in Proposition B.4 in the paper. When the correlation is positive but strictly less than one, the IV still has a downward bias but smaller than the OLS estimate.

We can test the prediction that the point estimates for $\delta$ should be larger when we instrument import penetration in the US using import penetration in other high-income countries than when we do not instrument. To do that, we re-run the main regressions in the paper without instrumenting for import penetration. Table B.2 shows side by side the OLS estimates and the IV estimates for the main outcome variables analyzed in the paper: sales growth, ROA, capital expenditures, and employment growth. The IV estimates are the same as in the paper. The OLS estimates are obtained by running the same regressions and replacing instrumented import penetration by actual China's import penetration in the US on the right-hand side (both for non-interacted import penetration and for import penetration interacted with R&D). For all four dependent variables, the OLS estimates have the same sign as the IV estimate. Consistent with the prediction of the model, the OLS estimates are smaller than IV estimates.

**Proof of Proposition B.6** We follow the same steps as in the proof of Proposition B.4 in the case $Corr(\mu, \mu^*) \neq 0$. In this case, we calculate

$$\begin{aligned}
Cov(z, R) &= -\frac{1}{\rho}\sigma_z^2, \\
Cov(z, M) &= 0, \\
Cov(z, RM) &= \frac{\lambda(\bar{\alpha} + \bar{\Omega})}{\rho}\sigma_z^2, \\
Cov(zM^*, R) &= \frac{\lambda(\bar{\alpha} + \bar{\Omega})}{\rho}\sigma_z^2, \\
Cov(zM^*, M) &= 0, \\
Cov(zM^*, RM) &= -\frac{1}{\rho}\Big(\lambda^2\big((\bar{\alpha} + \bar{\Omega})^2 + \bar{\Omega}^2 Cor(\mu, \mu^*)\sigma_\mu^2\big) + \sigma_T^2\Big)\sigma_z^2, \\
Cov(Y, z) &= -\frac{\bar{\gamma}}{\rho}\sigma_z^2, \\
Cov(Y, zM^*) &= -\frac{1}{\rho}\Big( - \bar{\gamma}\lambda(\bar{\alpha} + \bar{\Omega}) + \delta\sigma_T^2 - \bar{\gamma}\lambda\bar{\Omega}Cor(\mu, \mu^*)\sigma_\mu^2\Big)\sigma_z^2.
\end{aligned}$$

We calculate $Cov(\mathbf{Z}', \mathbf{M})^{-1}Cov(\mathbf{Z}', Y)$ and obtain:

$$E\Big[\hat{\delta}^{import-iv}\Big] = \delta - \frac{\Omega(\bar{\gamma} + \delta\bar{\Omega}\lambda)}{\sigma_T^2 + \bar{\Omega}^2\lambda^2 Cor(\mu, \mu^*)\sigma_\mu^2}\lambda\, Cor(\mu, \mu^*)\sigma_\mu^2 + o(1). \tag{B.14}$$

Table B.2: Without Instrumenting for Import Penetration

| | Sales growth | | ROA | | Capital expenditures | | Employment growth | |
|---|---|---|---|---|---|---|---|---|
| | OLS (1) | IV (2) | OLS (3) | IV (4) | OLS (5) | IV (6) | OLS (7) | IV (8) |
| Import penetration × R&D stock | 0.89** | 1.10*** | 1.29** | 1.42*** | 1.34** | 1.77*** | 0.75** | 0.91** |
| | (0.36) | (0.40) | (0.50) | (0.54) | (0.64) | (0.68) | (0.35) | (0.38) |
| Assets | 0.03*** | 0.03*** | 0.01 | 0.01 | 0.05*** | 0.05*** | 0.07*** | 0.07*** |
| | (0.01) | (0.01) | (0.01) | (0.01) | (0.02) | (0.02) | (0.01) | (0.01) |
| Age | -0.23*** | -0.23*** | 0.15*** | 0.15*** | -0.44*** | -0.44*** | -0.25*** | -0.25*** |
| | (0.02) | (0.02) | (0.02) | (0.02) | (0.03) | (0.03) | (0.02) | (0.02) |
| R&D stock | 0.07*** | 0.07*** | -0.21*** | -0.21*** | 0.03 | 0.02 | 0.14*** | 0.13*** |
| | (0.03) | (0.03) | (0.03) | (0.03) | (0.04) | (0.04) | (0.02) | (0.02) |
| Import penetration × Age | -0.50 | -0.67 | -0.27 | -0.22 | -0.02 | -0.34 | -0.53 | -0.82* |
| | (0.42) | (0.46) | (0.28) | (0.35) | (0.73) | (0.82) | (0.43) | (0.48) |
| Firm FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Industry-Year FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 23,907 | 23,907 | 24,533 | 24,533 | 24,321 | 24,321 | 23,197 | 23,197 |

The sample is US manufacturing firms over 1991–2007 from Compustat. Odd-numbered columns reproduce columns (4) of Tables VI to XII where we instrument import penetration in the US using import penetration in a group of eight other high-income countries (Australia, Denmark, Finland, Germany, Japan, New Zealand, Spain, and Switzerland). In even-numbered columns, we estimate the same regressions using directly (non-instrumented) import penetration in the US as a regressor. Standard errors are bootstrapped within industry-year clusters and reported in parentheses. *, **, and *** mean statistically different from zero at 10, 5, and 1% levels of significance.

# C Additional Empirical Analysis and Robustness Tests

## C.1 Import Penetration Scaled by Ten-Year Lagged Employment

In the paper, we scale import penetration using employment at the beginning of the sample period (in 1990). In this appendix, we show that our main results are robust to scaling import penetration using employment ten years before the beginning of the sample period (in 1980). Table C.1 reports the regression results of our preferred specification with industry-year fixed effects for the four main dependent variables used in the paper: sales growth, returns on asset, capital expenditures, and employment growth.

Table C.1: Import Penetration Scaled by Employment in 1980

| | Sales growth (1) | ROA (2) | Capital expenditures (3) | Employment growth (4) |
|---|---|---|---|---|
| Import penetration/(Employment in 1980 ) | 1.14*** | 1.57*** | 2.03*** | 0.93** |
| × R&D stock | (0.39) | (0.53) | (0.62) | (0.38) |
| Assets | 0.03*** | 0.01 | 0.05*** | 0.07*** |
| | (0.01) | (0.01) | (0.02) | (0.01) |
| Age | -0.23*** | 0.15*** | -0.43*** | -0.24*** |
| | (0.02) | (0.02) | (0.03) | (0.01) |
| R&D stock | 0.07*** | -0.21*** | 0.02 | 0.13*** |
| | (0.02) | (0.03) | (0.04) | (0.02) |
| Firm FE | Yes | Yes | Yes | Yes |
| Industry-Year FE | Yes | Yes | Yes | Yes |
| Observations | 23,333 | 23,967 | 23,756 | 22,591 |
| R2 | .34 | .71 | .39 | .31 |

The sample is US manufacturing firms over 1991–2007 from Compustat. We estimate the same regression as in column (4) of Tables VI to XII by scaling import penetration by industry employment in 1980 (instead of 1990 in the rest of the paper). Standard errors are bootstrapped within industry-year clusters and reported in parentheses. *, **, and *** mean statistically different from zero at 10, 5, and 1% levels of significance.

## C.2 Exogeneity of R&D Policy

A potential concern is that changes to the R&D tax credit policy may be endogenous. For example, do states offer more generous tax credits when they anticipate an increase in R&D expenditures? This issue is standard and have been been discussed in previous literature. For instance, Bloom, Schankerman and Van Reenen (2013) review the literature on US state R&D and corporate tax rates and conclude that "the existing literature suggests a large degree of randomness regarding the introduction and level of R&D tax credits" (page 1366). Papers that have tried to explain the evolution of state-level corporate tax credits have found that aggregate variables (such as the federal credit rate) have some explanatory power, but local economic or political variables do not seem important (e.g., Chirinko and Wilson (2008, 2011)). To investigate further that issue, we ask whether changes in R&D or economic activity predict changes in the tax-induced user cost of R&D capital that we use as our instrument.

We use three explanatory variables to predict changes in the R&D tax cedit. The first one is the change in state GDP measured over various horizons: from year $t-1$ to year $t$, from year $t-3$ to year $t$, and from year $t-5$ to year $t$. The second variable is the change in state-level R&D. To construct this variable, we compute, for each state-year, the weighted average ratio of R&D expenditures to total assets across all firms with inventors in the state, where firms are weighted by the share of their inventors located in the state (using the same weights $w_{ist}$ used in Section II.B to compute the weighted average user cost of R&D). Again, we consider changes in state R&D activity over the past 1, 3, and 5 years. The third variable is the change in the number of doctorates awarded in the state over the past 1, 3, or 5 years, which we obtain from the NSF WebCaspar database. We ask whether these variables predict the one-year ahead change in R&D tax credit $(\rho_{s,t+1} - \rho_{s,t})$. Results in Table C.2 show that, at all horizons, past changes in GDP, R&D, or number of doctorates do not predict changes in the R&D tax credit policy.

Table C.2: Exogeneity of R&D Tax Credit Policy

| | Change in state R&D tax credit ($t \rightarrow t+1$) | | |
|---|---|---|---|
| | (1) | (2) | (3) |
| | $h=1$ | $h=3$ | $h=5$ |
| Change in GDP ($t-h \rightarrow t$) | -.0022 | .0013 | .0053 |
| | (.0068) | (.0052) | (.0057) |
| Change in R&D ($t-h \rightarrow t$) | .0016 | -.00068 | -.0033 |
| | (.0057) | (.0022) | (.0031) |
| Change in number of doctorates ($t-h \rightarrow t$) | -.0023 | .0012 | .0011 |
| | (.0029) | (.0028) | (.0025) |
| State FE | Yes | Yes | Yes |
| Year FE | Yes | Yes | Yes |
| Observations | 1,224 | 1,224 | 1,224 |
| R2 | .92 | .92 | .92 |

The sample is a balanced panel over 51 US states and the years 1982 to 2007. We estimate a linear regression model where the dependent variable is the one-year ahead change in the tax-induced user cost of R&D. All regressions include state and year fixed effects. The regressors are the change in state GDP, the change in state R&D, and the change in the number of doctorates awarded in the state. The changes in the regressors are measured over the past year in column (1), over the past three years in column (2), and over the past five years in column (5). Standard errors are clustered by state and year and reported in parenthesis. *, **, and *** mean statistically different from zero at 10, 5, and 1% levels of significance.

## C.3 R&D Tax Incentives in Firms' Headquarter States

In the paper, we compute firm-level exposure to state tax credit based on the locations on their inventors. In this appendix, we show the robustness of our main results to measure firms' exposure to the tax credit based on the location of their headquarters. Results are reported in Table C.3.

Table C.3: R&D Tax Incentives in Firms' Headquarter States

| | Sales growth (1) | ROA (2) | Capital expenditures (3) | Employment growth (4) |
|---|---|---|---|---|
| Import Penetration × R&D stock | 1.19** | 1.59** | 2.30*** | 1.09** |
| | (0.48) | (0.66) | (0.82) | (0.46) |
| Import Penetration | -1.81 | -2.62*** | -1.06 | -0.53 |
| | (1.19) | (0.64) | (1.15) | (0.81) |
| Assets | 0.03*** | 0.01 | 0.06*** | 0.06*** |
| | (0.01) | (0.01) | (0.02) | (0.01) |
| Age | -0.22*** | 0.14*** | -0.44*** | -0.23*** |
| | (0.01) | (0.02) | (0.03) | (0.01) |
| R&D stock | 0.08*** | -0.26*** | 0.05 | 0.15*** |
| | (0.03) | (0.03) | (0.04) | (0.02) |
| Firm FE | Yes | Yes | Yes | Yes |
| Industry-Year FE | Yes | Yes | Yes | Yes |
| Observations | 23,321 | 23,959 | 23,748 | 22,576 |
| R2 | .32 | .75 | .34 | .39 |

The sample is US manufacturing firms over 1991–2007 from Compustat. We estimate the same regression as in column (4) of Tables VI to XII by measuring firms' exposure to the tax credit based on the location of their headquarters. Standard errors are bootstrapped within industry-year clusters and reported in parentheses. *, **, and *** mean statistically different from zero at 10, 5, and 1% levels of significance.

## C.4 Multi-Segment Firms

We use in the paper the historical main four-digit SIC industry to measure firms' exposure to import competition from China. This measure is noisy in the case of multi-segment firms since these firms can have operations in industries that are differently exposed to trade shocks. To refine our measure of exposure to Chinese import penetration, we use Compustat Business Segments data. These data provide disaggregated financial information for business segments that represent at least 10% of the firm's sales, assets, or profits.[7] 55% of firms in our sample report more than one business segment. We compute for each firm $i$ in each year $t$ the fraction of sales in each segment $j$ defined at the four-digit SIC code level: $f_{itj}$. We then construct import penetration at the firm-year level as the average of predicted Chinese import penetration in the US across all segments weighted by the share of each segment: $\sum_j f_{itj} ImportPenetration_{jt}$. Some multi-segment firms whose main SIC is in the manufacturing sector have operations in segments outside the manufacturing sector. Since the data for Chinese import penetration only cover manufacturing industries, part of these multi-segment firms' sales cannot be matched with the import penetration measure. We assume that segments outside the manufacturing sector are not exposed to import competition from China and assign a value of zero to import penetration for non-manufacturing industries. However, when a firm has more than 25% of its sales that cannot be matched with industry import penetration, we drop the observation, which excludes 10% of observations.

We re-run the second stage regression (6) of the paper using predicted import penetration based on segment sales. We adopt again our preferred specification with the full set of fixed effects and controls and use the same dependent variables as in the main analysis. Results are reported in Table C.4 and can be compared to columns (4) in Tables VI to XII in the paper. The estimated effects are qualitatively similar when using the main industry and when using segment-based weighted industries to construct the predicted value of import penetration. Depending on the dependent variable, coefficient estimates using the main SIC code range from one-fifth to one-third smaller than when using business

[7] These data are not without flaws. Villalonga (2004) documents that firms sometimes change the segments they report when there is no real change in their operations. This should not, however, affect the sign of the estimated effects to the extent that it only adds noise in the import penetration variable.

segments. This difference can reflect the fact that segment-based weighted industries are a more accurate proxy of a firm's true industry composition than the firm's main industry. The bias towards zero induced by noisy explanatory variables may thus be reduced in this case. In conclusion, our results are robust to, and even slightly strengthened by the use of business segment data to identify the industries of multi-segment firms.

Table C.4: Measuring Import Penetration From Business Segments

| | Sales growth | ROA | Capital expenditures | Employment growth |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Import penetration × R&D Stock | 1.51*** | 1.73** | 2.56*** | 1.18** |
| | (0.56) | (0.72) | (0.82) | (0.47) |
| Assets | 0.03*** | 0.01 | 0.05*** | 0.07*** |
| | (0.01) | (0.01) | (0.02) | (0.01) |
| Age | -0.22*** | 0.16*** | -0.43*** | -0.24*** |
| | (0.02) | (0.02) | (0.03) | (0.02) |
| R&D Stock | 0.06** | -0.23*** | 0.02 | 0.13*** |
| | (0.03) | (0.03) | (0.04) | (0.02) |
| Import penetration × Age | -0.03 | 0.20 | -0.04 | -0.34 |
| | (0.51) | (0.49) | (0.86) | (0.54) |
| Firm FE | Yes | Yes | Yes | Yes |
| Industry-Year FE | Yes | Yes | Yes | Yes |
| Observations | 22,271 | 22,888 | 22,699 | 21,620 |
| R2 | .33 | .72 | .4 | .35 |

The sample is US manufacturing firms over 1991–2007 from Compustat. We estimate the same regression as in column (4) of Tables VI to XII except that we now identify firms' industries using Compustat business segments. The segment-based predicted import penetration variable is computed as the average predicted import penetration across all the segments of the firm weighted by the share of each segment. The coefficient on import penetration is absorbed by industry-year fixed effects. Standard errors are bootstrapped within industry-year clusters and reported in parentheses. *, **, and *** mean statistically different from zero at 10, 5, and 1% levels of significance.

## C.5 Excluding California

27% of firms in our sample have more than half of their R&D activity in California. To check that our results are not driven by California, we re-run our regressions on sales growth, ROA, capital expenditures and employment growth after excluding these firms. We use our preferred specification with the full set of controls and fixed effects, as in columns (4) of Tables VI to XII. Results are reported in Table C.5. They are similar to the results on the entire sample both in terms of statistical significance and economic magnitude. The point estimates are similar for the effect on sales and profitability and somewhat larger for the effect on capital expenditures and employment. Standard errors are a bit higher than when using the entire sample, which is due to the fact that the number of observations drops by 27%.[8] Overall, California does not drive our results.

[8] Reducing the sample size by 27% mechanically increases the standard error by $1/\sqrt{1-27\%}-1 = 17\%$.

Table C.5: Robustness: Excluding Californian Firms

| | Sales growth | ROA | Capital expenditures | Employment growth |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Import penetration × R&D stock | 1.21** | 1.41** | 2.77*** | 1.75*** |
| | (0.58) | (0.62) | (0.89) | (0.54) |
| Assets | 0.03*** | 0.01 | 0.05*** | 0.06*** |
| | (0.01) | (0.01) | (0.02) | (0.01) |
| Age | -0.20*** | 0.13*** | -0.38*** | -0.23*** |
| | (0.02) | (0.02) | (0.03) | (0.02) |
| R&D stock | 0.08** | -0.19*** | 0.00 | 0.11*** |
| | (0.03) | (0.03) | (0.04) | (0.02) |
| Import penetration × Age | -0.48 | -0.24 | -0.13 | -0.49 |
| | (0.55) | (0.45) | (0.77) | (0.44) |
| Firm FE | Yes | Yes | Yes | Yes |
| Industry-Year FE | Yes | Yes | Yes | Yes |
| Observations | 17,471 | 17,876 | 17,712 | 16,970 |
| R2 | .35 | .73 | .42 | .37 |

The sample is US manufacturing firms over 1991–2007 from Compustat, excluding firms that have more than 50% of their investors located in California. We estimate the same regression as in column (4) of Tables VI to XII. Standard errors are bootstrapped within industry-year clusters and reported in parentheses. *, **, and *** mean statistically different from zero at 10, 5, and 1% levels of significance.

## C.6 Test for Non-Monotonic Effect of Import Competition

The results in Section III show that an increase in import competition leads to an *average* increase in the performance differential between innovative firms and less innovative firms. Aghion et al. (2005) suggest that this relation may be inverse U-shaped. To test whether the average positive relation we uncover in the data conceals a non-monotonic relation, we estimate a specification where the return from R&D is a quadratic function of import competition. To do so, we interact the predicted R&D stock with both predicted import penetration and its square.[9] We adopt our preferred specification with firm and industry-year fixed effects as well as firm age interacted with import penetration and squared import penetration to control for the correlation between firm age and instrumented R&D stock, as in columns (4) of Tables VI to XII.

In Table C.6, we estimate this quadratic specification on the dependent variables we have considered in Sections III and V: sales growth, ROA, capital expenditures, and employment growth. The interaction between R&D and imports is positive and significant while the interaction between R&D and squared imports is negative (and significant for three out of four variables and the fourth one has a $p$-value of 0.14). The point estimates imply that the returns to R&D become decreasing in import competition when import penetration goes above 100 k$/worker, which corresponds to the 99th percentile of the sample distribution. Thus, the relation between import competition and return to R&D is positive almost everywhere.

[9] We have also done this analysis using the linear prediction of squared import penetration (instead of taking the square of the linear prediction of import penetration), where in the first stage we predict both import penetration in the US and squared import penetration in the US using import penetration in the other 8 high-income countries and squared import penetration in the other 8 high-income countries. Results are very similar with this alternative specification.

Table C.6: R&D Capital in Import-Competing Industries: Quadratic Specification

| | Sales growth | ROA | Capital expenditures | Employment growth |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Import penetration × R&D stock | 2.63** | 2.72** | 6.82*** | 3.09*** |
| | (1.07) | (1.17) | (1.75) | (0.90) |
| Import penetration$^2$ × R&D stock | -13.13* | -11.61 | -43.76*** | -18.76*** |
| | (7.75) | (7.87) | (12.78) | (6.71) |
| Assets | 0.03*** | 0.01 | 0.05*** | 0.07*** |
| | (0.01) | (0.01) | (0.02) | (0.01) |
| Age | -0.23*** | 0.15*** | -0.44*** | -0.24*** |
| | (0.02) | (0.02) | (0.03) | (0.02) |
| R&D stock | 0.07** | -0.22*** | 0.00 | 0.12*** |
| | (0.03) | (0.03) | (0.04) | (0.02) |
| Import penetration × Age | -0.76** | -0.63*** | -1.13** | -0.55** |
| | (0.30) | (0.24) | (0.51) | (0.27) |
| Import penetration$^2$ × Age | 1.53 | 1.14 | 5.77** | 1.80 |
| | (1.45) | (1.20) | (2.42) | (1.26) |
| Firm FE | Yes | Yes | Yes | Yes |
| Industry-Year FE | Yes | Yes | Yes | Yes |
| Observations | 23,907 | 24,533 | 24,321 | 23,197 |
| R2 | .34 | .72 | .41 | .36 |

The sample is US manufacturing firms over 1991–2007 from Compustat. We estimate the same regression as in column (4) of Tables VI to XII with as additional regressors squared predicted China import penetration interacted with predicted stock of R&D capital, and predicted China import penetration interacted with log of firm age. Standard errors are bootstrapped within industry-year clusters and reported in parentheses. *, **, and *** mean statistically different from zero at 10, 5, and 1% levels of significance.

## C.7 Other Proxy for Differentiation from Chinese Competitors

In the paper, we show that firms with a higher level of (exogenous) R&D increases differentiation following import competition shocks by proxying for differentiation from Chinese Competitors with differentiation from US competitors. In this appendix, we present the results using another proxy designed to capture more precisely differentiation from Chinese competitors. The idea is to test whether a firm's products become more similar to the products of other US firms less exposed to Chinese competition when the firm has a higher stock of R&D. For each firm $i$ in each year $t$, we compute the weighted average import penetration faced by the firms to which firm $i$ is close in the product market space:

$$ChineseProductsSimilarity_{it} = \sum_{k \neq i} w_{ikt} \times ImportPenetration_{kt}, \tag{C.1}$$

where the weight $w_{ikt}$ on firm $k$ is proportional to the Hoberg and Phillips (2015) product similarity index between firm $i$ and firm $k$:

$$w_{ikt} = \frac{ProductSimilary_{ikt}}{\sum_{k' \neq i} ProductSimilary_{ik't}}, \tag{C.2}$$

and $ImportPenetration_{kt}$ is the instrumented measure of import penetration of firm $k$. The value of $ChineseProductsSimilarity_{it}$ is large when firm $i$'s products are similar to the products of US firms exposed to competition from Chinese firms; it is low when firm $i$'s products are more similar to the products of US firms not exposed to Chinese competition. Thus, a lower value of this variable reflects more differentiation from Chinese firms. We estimate the same specification as before using $ChineseProductsSimilarity$ as the dependent variable.

Results are reported in Table C.7. In column (1) the positive coefficient on import penetration is purely mechanical: when a firm is more exposed to China's import penetration (higher value of the RHS variable), it is closer in the product market space to other US firms that are also exposed to China's import penetration (higher value of the LHS variable). The more interesting question is whether this relation is weaker for firms with a higher stock of R&D. In column (2) we interact import penetration with R&D and obtain a negative and statistically significant coefficient on the interaction term. Thus, when competition from China increases, firms with a higher level of R&D are able to differentiate their products in a way that makes their products less similar to the products of US firms in import-competing industries and more similar to the products of US firms less exposed to Chinese competition. When we include industry-year fixed effects, the point estimate barely changes but the statistical significance is weaker ($p$-value is 0.20, columns (3) and (4)). Overall, the results using this other proxy point in the same direction as those reported in the paper: R&D allows firms to escape import competition by improving their ability to differente when the competitive pressure increases.

Table C.7: Other Proxy for Differentiation from Chinese Competitors

| | Chinese products similarity | | | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Import penetration | .3*** | .33*** | | |
| | (.028) | (.031) | | |
| Import penetration × R&D stock | | -.064** | -.039 | -.041 |
| | | (.029) | (.031) | (.031) |
| Assets | -.00019 | -.00034 | -.0002 | -.0002 |
| | (.00019) | (.00025) | (.00021) | (.00021) |
| Age | -.000079 | .000043 | .00018 | .00065 |
| | (.00054) | (.00057) | (.00052) | (.00055) |
| R&D stock | | .00042 | .00033 | .00036 |
| | | (.0006) | (.00053) | (.00053) |
| Import penetration × Age | | | | .05 |
| | | | | (.035) |
| Observations | 16,061 | 16,061 | 16,061 | 16,061 |
| Firm FE | Yes | Yes | Yes | Yes |
| Year FE | Yes | Yes | – | – |
| Industry-Year FE | No | No | Yes | Yes |

The sample is US manufacturing firms over 1996–2011 from Compustat. We estimate a linear regression model on a firm-year panel where the dependent variable is defined in equations (C.1)–(C.2) and is equal to the weighted average China's import penetration of the firm's peers where the weights are proportional to the Hoberg and Phillips (2015) product similarity index with respect to the peers. All specifications include firm fixed effects, year fixed effects in columns (1) and (2), industry-by-year fixed effects in columns (3) and (4), and log of total assets and log of firm age as controls. $ImportPenetration$ is industry-year-level import penetration from China in the US instrumented using China import penetration in eight other high-income markets. $R\&DStock$ is firm-year-level predicted stock of R&D capital instrumented using firm-specific tax-induced user cost of R&D capital. Standard errors are bootstrapped within industry-year clusters and reported in parentheses. *, **, and *** mean statistically different from zero at 10, 5, and 1% levels of significance.

# References


Aghion, P., Bloom, N., Blundell, R., Griffith, R., and Howitt, P., 2005. "Competition and Innovation: an Inverted-U Relationship". *Quarterly Journal of Economics* 120, 701-728.

Aghion, P., and P. Howitt, 1992. "A Model of Growth Through Creative Destruction." *Econometrica* 60(2), 323-354.

Bloom, N., Schankerman, M., and Van Reenen, J., 2013. "Identifying Technology Spillovers and Product Market Rivalry". *Econometrica* 81, 1347-1393.

Chirinko, R.S., and Wilson, D.J., 2008. "State Investment Tax Incentives: A Zero-Sum Game?" *Journal of Public Economic* 92, 2362-2384.

Chirinko, R.S., and Wilson, D.J., 2011. "Tax Competition Among U.S. States: Racing to the Bottom or Riding on a Seesaw". *Working Paper.*

Gabaix, X, 2011, "The Granular Origins of Aggregate Fluctuations." *Econometrica* 79, 733-772.

Grossman, G. and Helpman, E., 1991. "Quality Ladders in the Theory of Growth" *Review of Economic Studies* 58(1), 43-61.

Hoberg, G., and G. Phillips 2015. "Text-Based Network Industries and Endogenous Product Differentiation." *Journal of Political Economy, Forthcoming.*

Sutton, John. 1991. "Sunk Costs and Market Structure" *(MIT Press: Cambridge, Mass).*

Tirole, J., 1988. *The theory of industrial organization.* MIT press.

Villalonga, B., 2004. "Diversification Discount or Premium? New Evidence from the Business Information Tracking Series". *Journal of Finance* 59, 479-506.