

# Do Imports Spur Incremental Innovation in the South?

Yi Lu\* and Travis Ng†

March 2012

## Abstract

We estimate that a one-standard-deviation increase in a firm's import penetration ratio raises its likelihood of having engaged in an incremental innovation by 4.48% using a random-sampled firm survey in China. The estimate is close to those in Gorodnichenko, Svejnar and Terrell (2010). A number of empirical strategies rule out alternative explanations as sufficient drivers of our result. Competitive pressure from imports is shown to be an underlying mechanism through which imports spur incremental innovation. We discuss how the link between imports and innovation in the South differ from that in the North.

**Keywords:** Imports, Import Competition, Trade, Innovation, Incremental Innovation

**JEL classification:** F10, F14, O31, L20.

---

\*Department of Economics, The National University of Singapore. Email: ecluyi@nus.edu.sg

†Department of Economics, The Chinese University of Hong Kong. Email: TravisNg@cuhk.edu.hk.

# 1 Introduction

This paper estimates that a one-standard-deviation increase in the import penetration ratio a firm faces raises its likelihood of having engaged in an incremental innovation by 4.48%. We use a random-sample of firms surveyed by the World Bank in 2002 in China. A firm is referred to as having engaged in incremental innovation if it has introduced either new products/services, business lines, management practices, quality controls, or production processes in the past three years.<sup>1</sup> Unlike the traditional use of industry-level import penetration ratio, our firm-level import penetration measure equals the share of imports of a firm's major market. Our 4.48% estimate is in the neighborhood of those reported in Gorodnichenko, Svejnar and Terrell (2010); using firms surveyed in European developing countries, their estimates are 7% and 4% for new product and new technology, respectively.

Our result is important in three aspects. First, developing countries have been increasingly innovative; investigating its underlying drivers is important. Second, developing countries have become significantly more globalized. Linking their globalization trend with their innovation may draw potentially important policy implications. Third, Leamer (2007) argues that the unintended transfer of innovative processes in manufacturing, alongside the outsourcing of production, tops all other concerns of "the world is flat." The developed world should thus examine the causes and consequences of these conditions. Indeed, developing countries are not only *exporting* more to the rest of the world, they are also *importing* more from the rest of the world. Do their exports to the developing countries also spur innovativeness there?

The strong association between import penetration and incremental innovation can be due to the possibility that more innovative firms are also more globally-connected

---

<sup>1</sup>The literature refers "process innovations" as small-step cost-cutting innovations. In contrast, "incremental innovations" embody more than just innovations on processes; such innovations include product innovations, new quality controls, new management practices, and so on. Tailoring existing technology to niche markets at a large volume is an example: Haier gains a 60% share of the world's wine-storage fridges, a niche market in fridge manufacturing.

through foreign ownership and international trade. The strong association may just reflect such connections. We rule out this alternative explanation by directly controlling for foreign ownership, fraction of sales abroad, and fraction of imported inputs.

Another alternative explanation is that more capable firms innovate more and are more likely to compete in markets with intense import competition. We use four strategies to show that this alternative explanation is unlikely to drive our results.

Our first strategy is to directly control for firm and CEO characteristics, most notably, lagged labor productivity, levels of capital, and employment size. We also include CEOs' education level and political capital.

Our second strategy is to address omitted variable bias by instrumenting our regressor of interest, firm-level import penetration ratio, with an instrumental variable (IV) that is unlikely to correlate with any omitted firm-level variables. Our IV is the average of the import penetration ratios faced by firms belonging to the same industry but located in different cities. Section 4.1 details the relevance of this IV and why it is unlikely to fail the exclusion restriction.

Our third strategy is the use of the imperfect IV method developed in Nevo and Rosen (2011). The idea is to relax the IV's identifying assumption by assuming it to be endogenous, but less so than our regressor of interest. Relaxing the identifying assumption does not remove the significant association between import penetration ratio and incremental innovation.

Our fourth strategy is to address a plausible explanation that technologically capable firms may self-select to locate at markets with intense import competition. We fail to document compelling evidence supporting this plausible explanation.

The four strategies, along with a battery of other robustness checks using alternative measures and alternative specifications, and controlling for sample attrition bias, suggest that the alternative explanations and other econometric issues are unlikely to drive the strong association between import penetration ratio and incremental innovation.

Section 5 examines indirectly whether competition effect (i.e., foreign imports force domestic firms to improve their production) is a channel through which imports affect firms' incremental innovation. If such is the case, the firms that have pre-innovation rents less affected by imports should have a weaker correlation between their innovation and imports. Section 5.1 uses the heterogeneous response estimation à la Rajan and Zingales (1998) to document such a pattern. Firms with a larger fraction of sales to governmental agencies, which implies that their pre-innovation rents are less affected by imports, have their innovation correlated significantly less with imports.

We also show evidence consistent with the discouragement effect (i.e., domestic firms give up competing with imports because foreign rivals are significantly more superior) in Aghion, Bloom, Blundell, Griffith and Howitt (2005). Firms in mature manufacturing industries with products relatively closer to global technological frontiers respond to imports to innovate, but not firms in high-tech manufacturing industries with products relatively far away from global technological frontiers.

Our paper echoes Gorodnichenko et al. (2010). They document that firms in European developing countries also respond to imports by engaging in more small-step innovations. We supplement their results by identifying a similar response among Chinese firms. China is important in examining the link between imports and innovation in the South. First, China's innovation is significant. Between the period from 1990 to 1992 and 2000 to 2002, China's share of innovative new goods in US imports increased from 17% to 48% or from 0% to 5% in terms of trade value. Second, it spends the most in R&D among developing countries. China's expenditure on R&D reached USD 86.8 billion in 2006 after an impressive growth of approximately 19% annually in real terms from 2001 to 2006 (OECD, 2008). China claimed 2% of the global share of total R&D expenditure in 1996, dramatically increasing to 7.5% in 2005, just behind the US (35%), the EU27 (24%), and Japan (14%). Third, while a world exporter, China is also a gigantic importer. The WTO (2005) ranks China as the world's third largest importer of merchandises in 2004

(USD 561.4 billion or 5.0% of the world's total) and eighth largest importers of services (USD 69.7 billion or 3.3% of the world's total).

Teshima (2009) measures import competition faced by Mexican firms based on firm-specific import tariff reduction. He shows that tariff reduction raises spending on both process R&D and product R&D. While his main focus is R&D spending, ours is on innovation output rather than input. His import competition measure is also different from ours. To the extent that more R&D spending produces more innovation, our result is closely in-line with his.

## **2 Linking imports with incremental innovation in the South**

Aghion et al. (2005) show competition can either promote or discourage innovation. If increased competition reduces pre-innovation rents more than post-innovation rents, firms "escape competition" by innovating (the competitive effect). On the other hand, if pre-innovation rents are low to begin with, increased competition primarily reduces post-innovation rents. Firms are thus discouraged from innovating (the discouraging effect). The competitive effect likely dominates if a firm has technology on par with that of its competitors. In contrast, the discouraging effect likely dominates if a firm is laggard with low pre-innovation rents to begin with.

Unlike firms in the North, firms in the South are more likely to lag behind their import competitors because North-North trade and North-South trade dominates South-South trade (OECD, 2006; UNCTAD, 2005). The OECD (2006) estimates merchandise exports and trade in services among countries in the South in 2002 to be only 6% and 10%, respectively, of the corresponding world figures. The discouraging effect is therefore likely to dominate in the South.

The effect of increased competition on innovation also depends on the beginning level of competition. If the degree of competition is low to begin with, the competition effect

dominates, and vice versa. Although the South does import from the rest of the world, its trade barriers are considerably higher than those of the North. The OECD (2006) estimates that in 2001 the North faced average tariff rates of 4.4 and 5.0 for imports from other countries in the North and imports from the South, respectively. The corresponding rates are 9.9 and 11.1, respectively, in the South. They also liberalize their trade later than the North. The South is thus relatively more likely to begin with a lower level of import competition than the North. The competition effect may dominate instead.

While Aghion et al. (2005) proxy innovation with patenting, Gorodnichenko et al. (2010) note that patenting cannot be a desirable measure of innovation in the South. First, the propensity to patent an innovation varies across countries and regions. Second, methods other than patenting in protecting intellectual properties are likely more effective because the South has weaker formal institutions. Moreover, incremental innovation involves small-step innovation, which are less likely to justify the legal fee of patenting.

Incremental innovation of the South can also be relevant to the exogenous parameter  $h$  (the rate at which “laggard firms” move one-step forward along the technology line by copying others’ technology) in Aghion et al. (2005). Increased imports from more advanced countries can strengthen this copying effect (increased imports may raise  $h$ ) because reverse-engineering or imitating an innovation of a rival is easier than developing an innovation de novo. Schott (2008) shows that China’s exports to the US overlap to a surprising extent with those from the OECD countries. Rodrik (2009) shows that China’s exports are sophisticated as those of a country three times richer. These evidence is consistent with a significant copying effect in the South.

Foreign imports may also change the taste and preferences of domestic consumers. Domestic firms may engage in innovation to respond to such changes.<sup>2</sup> For example, if a season is the old norm for fashion updates, Zara may have changed the norm in numerous countries into weeks.<sup>3</sup> To the extent that foreign imports fundamentally change the tastes

---

<sup>2</sup>We thank Steven J. Davis for suggesting this mechanism.

<sup>3</sup>“Fashion for the masses - Global stretch: When will Zara hit its limits?” *The Economist*, March 10, 2011.

and preferences of domestic consumers, firms in the South are likely to innovate more as a response to such changing preferences.

How increased imports affect incremental innovation in the South is ultimately an empirical question. The competitive, copying, and preference-changing effects suggest a positive relationship, whereas the discouraging effect implies a negative relationship. The nature of incremental innovation and the trade pattern of the South reflect, interestingly, some relevance of each of these effects in ways different from those in firms and industries in the North.

### 3 Data and Variables

We use data from the *Survey of Chinese Enterprises* (SCE) conducted by the World Bank in early 2003. The Data Appendix details the dataset. This section introduces the variables and summarizes their strengths and limitations.

#### 3.1 Incremental innovation

The SCE captures the concept of incremental innovation by asking whether the firm has introduced any new products/services, business lines, management practice, quality controls, and production processes over the past three years. The dummy *Innovation* takes the value of 1 if the firm answers yes to any of these questions, and 0 otherwise. Table 1 shows that slightly less than 70% of the firms have engaged in either one of the five types of incremental innovations.

Measurement errors may contaminate *Innovation*. First, the five categories cover some but not all possible types of innovations a firm could have taken. Second, so long as different firms perceived the term “new” differently, this self-reported variable is likely to be more noisy than an otherwise ideal measure of innovation that is objectively defined,

---

<http://www.economist.com/node/18333093>

and audited by researchers across firms. Conditional on our controls, if these extra noises are random, they would not bias the slope estimate but would inflate its standard error. In turn, a larger standard error works against us by decreasing the likelihood of finding a significant association between imports and innovation.

For robustness check, we follow Lin, Lin, and Song (2010) by setting a dummy (denoted by *R&D*) to 1 if the R&D spending in 2002 is positive, and 0 otherwise. This alternative measure is available for manufacturing firms only. Approximately 42% of the manufacturing firms had non-zero R&D spending. Similar to the measure in Teshima (2009), *R&D* differs from *Innovation* because it considers research input rather than output.

### **3.2 Imports**

We measure imports at the firm level. The variable *Import penetration* is the natural-log of (one plus the percentage of imports in the firm's major market). Table 1 indicates that, on average, firms face 6.43% of import penetration in their major markets. Gorodnichenko et al. (2010) use firms' responses to the perceived severity of foreign competition on a scale of one to four. Although our measure is different, our estimate is shown later to be close to their estimates.

Using firm-level, rather than industry-level import penetration ratio has several advantages. Using industry-level import penetration to proxy a firm's import threat entails a trade-off between using aggregated versus disaggregated industry classification. A more disaggregated industry-level measure of import penetration avoids the lumping of intermediate imports with processed imports, given that these two types of imports can affect firms' innovation differently.

On the other hand, a large number of firms engage in more than one industry. Grouping firms into industries can be arbitrary, more so under more disaggregated industry classification. Using firm-level import penetration ratio avoids such a trade-off. We also

do not have to assume that all firms within an industry are exposed to the same degree of foreign competition. Suppose firm A focuses on the inland regions, whereas firm B in the same industry focuses on the coastal regions. Assume that both firms face the same level of imports is problematic.

A cost of using our firm-level import penetration, however, is its inherent endogeneity. Firms may shift their major market in response to imports. In addition, the extent of a major market can also vary across firms. Some firms' major markets are national, whereas others' are regional instead. Firms with a few major markets may also report *Import penetration* by arbitrarily aggregating imports across their major markets. Finally, firms may mis-measure/mis-report imports.

Such potentially arbitrary aggregation and the mis-reporting make our measure more noisy than an otherwise ideal measure of imports that has an objectively defined market with researcher-audited import figures. In multiple regressions, these measurement errors may generate an attenuation bias only if the errors are truly random. However, the measurement errors may not be random and may reflect unobserved firm heterogeneity in innovation capability. If such is the case, these measurement errors can also create an upward bias in the estimate of *Import penetration* on *innovation*.<sup>4</sup> This issue is an important concern not to be dismissed.

We therefore also construct an alternative measure: *Importers* equals the natural-log of (one plus the percentage of overseas competitors), where the percentage is the answer to the question: "considering all your competitors, what percentage (in terms of output) have located their plants overseas?" Table 1 indicates that firms, on average, have 5.69% of their competitors being importers, which is in the neighborhood of the 6.43% import penetration. Although the two measures are consistent, we cannot entirely rule out measurement errors. Thus, we also rely on instrumental variable estimation for robustness checks.

---

<sup>4</sup>We are grateful for Steven J. Davis and a referee for pointing out this important concern.

Shifting major markets in response to imports is another possible source of endogeneity, particularly if we look at breakthrough innovations that are usually outcomes of long-term and major R&D investments. In contrast, incremental innovations are small-step innovations that are more optimization-based. We would expect that incremental innovations are more easily varied choices relative to shifting a firm's major markets, which requires long-term planning and marketing effort. Nonetheless, in our empirical analysis, we take this possible endogeneity seriously by controlling for a host of variables, and by using an IV and the technique of imperfect IV. Moreover, Section 4.2 uses the industry-city average of *Import penetration* as an alternative measure.

### 3.3 Other controls

To disentangle the effect of import competition on innovation from other globalization forces, we directly control for other global connections. As in Gorodnichenko et al. (2010), we control for *Overseas customers* (percentage of overseas customers), and *Overseas suppliers* (percentage of overseas suppliers). In addition, we control for *Foreign share* [natural-log of (one plus the percentage shares owned by foreign individuals, foreign institutional investors, foreign firms, and foreign banks in 1999)].

To rule out the alternative explanation that technologically-capable firms both innovate more and locate in markets with intense import competition, we directly control for a host of firm and CEO characteristics that proxy individual firms' technological capability. Firm characteristics include *Labor productivity* (natural-log of sales-over-employment in 2000), *Capital stock* (natural-log of total fixed assets in 2000), *Firm size* (natural-log of employment), *Firm age* (natural-log of years of establishment), *State ownership* (share owned by the government in 1999), and *Degree of computerization* (share of workers regularly use computer in their jobs), whereas institutional quality involves *Property rights protection* (percentage of government officials oriented toward helping rather than hindering firm operations) and *Contract enforcement* (likelihood that the legal system will uphold

the firm's contract and property rights in business disputes).

The CEO characteristics are measures of his/her human capital, including *CEO education* (years of schooling), *CEO tenure* (years as a CEO), and *Deputy CEO previously* (a dummy indicating whether the CEO was the deputy CEO before), as well as his/her political capital, which includes *Government cadre previously* (a dummy indicating whether the CEO was a government official before) and *Party member* (a dummy indicating whether the CEO was a member of the Chinese Communist Party).<sup>5</sup>

Bloom, Romer and Van Reenen (2010) stress the importance of accounting for labor and product market differences in identifying trade effects. We thus control for city and industry dummies. In addition, we separate foreign competition from local competition by including *Agglomeration* (percentage of output in the firm's major market supplied by competitors within the same city), as well as competition dummies corresponding to the question on the number of competitors of the firm's main business line in its major market.<sup>6</sup>

Section 4.1 uses the average of *Import penetration* among firms belonging to the same industry but located in other cities as the IV. Section 5.1 uses the heterogenous response estimation à la Rajan and Zingales (1998) to test whether the correlation of imports on innovations is weaker if firms' pre-innovation rents are less affected by imports. To proxy the degree of exposure to competition, we use the share of goods sold to government agencies (denoted by *State sales*).

## 4 Empirical Analysis

We estimate the following equation:

$$Innovation = \alpha + \beta \cdot Import\ penetration + X' \Gamma + \varepsilon, \quad (1)$$

---

<sup>5</sup>Li, Meng, Wang, and Zhou (2008) use these variables to examine the impact of political connections on business performance.

<sup>6</sup>Ayyagari, Demircuc-Kunt, and Maksimovic (2008) use the same measures.

where  $X$  is the set of controls. To deal with heteroskedasticity, the standard errors are White-corrected and are clustered at the industry-city level.

Panel A of Table 2 shows the OLS estimation results. Column 1 shows that *Import penetration* is associated positively with *Innovation*. Controlling for industry and city dummies in Column 2 does not change this result.

Columns 3 and 4 include firm and CEO characteristics. In particular, controlling directly for firms' global connection is an important strategy for isolating the effects of import competition on innovation from the other globalization forces. Larger, younger, more productive firms, as well as those that are less agglomerated, with more human capital (proxied by the average usage of computers among the workforce), and with more capable CEOs innovate more. Global connection is not significantly correlated with innovations. Column 5 controls for local competition to separate domestic from foreign competition. A significantly positive association between *Innovation* and *Import penetration* persists across all columns.

Panel B reports the corresponding marginal effects of *Import penetration* (evaluated at means) using Probit. A similar pattern emerges, suggesting that the positive association cannot be driven by the use of the linear probability model.

Experimenting different sets of controls in Columns 3 to 5 does little to the size of the estimate in both panels, suggesting that although ruling out omitted variable bias is impossible, omitted variables are unlikely to drive the positive association.

We also compare the economic significance of our estimate with that of Gorodnichenko et al. (2010). The standard deviation of *Import penetration* is 0.12, and the marginal effect from Probit is 0.259. The increase in probability as a percentage of the average of *Innovation* is  $\frac{0.12}{0.693} \cdot 0.259 = 4.48\%$ . Therefore, a one standard deviation increase in *Import penetration* corresponds to a 4.48% increase in the average probability of *Innovation*. This estimate is in the neighborhood of their estimates of 7% and 4%, respectively, for the increase in the probability of developing new product and new technology.

## 4.1 Instrumental variable estimation

Aside from omitted variable bias, endogeneity may result from potential reverse causality: markets more flooded with foreign imports might be a luxury that only more innovative firms can afford. In addition, as discussed, measurement errors may generate biases. Motivated by Hausman, Leonard and Zona (1994), we instrument *Import penetration* with the average of *Import penetration* among firms belonging to the same industry but located in other cities (denoted  $Imp_{i,-c}$ ).

Suppose the firm-level measure of *Import penetration* (denoted  $imp$ ) is related to the industry-city level, i.e.,

$$imp_{fic} = \gamma imp_{ic} + \eta_c + \lambda_i + v_{fic}, \quad (2)$$

where  $f$ ,  $i$ , and  $c$  index firm, industry and city, respectively. Then  $imp_{ic}$  becomes a natural IV for  $imp_{fic}$ : this industry-city average is unlikely to correlate with any omitted firm-level variables because of random sampling of firms. However, a variety of industry-city level shocks may affect industry-city-level innovation ( $Innovation_{ic}$ ) and firm-level innovation ( $Innovation_{fic}$ ). If these shocks are reflected by the industry-city-level import penetration ( $imp_{ic}$ ), the IV can be problematic.

We exploit the structure of the data by rewriting  $imp_{ic}$  as

$$imp_{ic} \equiv c_i \cdot imp_i - \sum_{c' \neq c} imp_{ic'}, \quad (3)$$

where  $imp_i$  is the industry-average import penetration;  $c_i$  is the number of cities. Hence,

$$\begin{aligned} imp_{fic} &= -\gamma \sum_{c' \neq c} imp_{ic'} + \gamma c_i \cdot imp_i + \eta_c + \lambda_i + v_{fic} \\ &= -\gamma imp_{i,-c} + \eta_c + \tilde{\lambda}_i + v_{fic} \end{aligned} \quad (4)$$

where  $c_i \cdot imp_i$  is absorbed by the industry dummy  $\tilde{\lambda}_i$ ; and  $imp_{i,-c} = \frac{1}{c_i-1} \sum_{c' \neq c} imp_{ic'}$  is our IV. Controlling for the industry and city dummies, our IV is negatively correlated

with  $imp_{fic}$ .

The intuition is that the amount of total imports reflects the industrial policies and the relevant transaction costs of international trade. For example, the more protective the government is of one industry, the lower is this industry's total imports. Melitz's (2003) model, for instance, predicts that falling trade cost results in shrinking domestic market share because some unproductive firms exit and all surviving plants sell less to the domestic market. Given that the industry dummies control for the industrial policies, the inter-city difference within an industry reflects the differential impacts of industrial policies across different cities. Conditioning on the total amount of imports, if the effect of a set of policies is larger in one city, it is also smaller in another city.

The identifying assumption is that conditional on our controls, the IV does not affect innovation through channels other than a firm's own import penetration ratio, i.e.,

$$E [imp_{i,-c} \cdot \varepsilon | X] = 0. \quad (5)$$

This assumption is plausible because the above-mentioned industry-city-level shocks that affect innovation are less likely to be picked up by the import penetration ratios faced by firms in other cities. While measurement errors can be potentially serious at the firm-level, the average of firms' import penetration likely averages out these errors. The IV is therefore unlikely to systematically correlate with any firm-level measurement errors.

Panel B of Table 3 shows that the IV is negatively correlated with the endogenous variable. The F-statistic is significantly above the critical value (10) of the "safe zone" for strong instrument (Staiger and Stock, 1997), which helps rule out the concern of weak instruments. Panel A shows that *Import penetration*, after being instrumented, still associates positively and significantly with *Innovation*. Panel C shows the corresponding marginal effect from IV Probit estimation.

The IV estimates are larger than the corresponding OLS estimates. The increase in size

is consistent with our concern that there can be potentially serious measurement errors on *Import penetration* that bias down the OLS estimates. However, our IV may also correlate positively with the error term. This failure of the orthogonal condition (5) will bias up the IV estimates. We address this concern in the following section.

#### 4.1.1 Imperfect IV

Nevo and Rosen (2011) innovate a technique wherein an imperfect IV is used to draw inference; the technique admits relaxing the orthogonal condition of the instrument (5), resulting in bounds of the parameter of interest. A considerably *weaker* assumption, that the IV may correlate with the error term in the same direction as the endogenous variable (Assumption 3 in Nevo and Rosen, 2011) but the correlation is weaker than that of the endogenous variable (Assumption 4 in Nevo and Rosen, 2011), is made.<sup>7</sup>

If more innovative firms tend to stay in markets with more imports, to the extent that our controls do not entirely rule out such possibilities, we would expect a positive correlation between *Import penetration* and the error term. The upward estimates of the IV estimation also may be an artifact of a positive correlation between the IV and the error term. Hence, Assumption 3, that our IV and the endogenous variable correlate with the error term in the same direction is plausible. In addition, given that our estimation is at the firm-level, arguing that our IV (industry-city level) is even more endogenous than our endogenous variable (firm-level) is implausible. Hence, Assumption 4 is likely to be true. We then apply Proposition 3 in Nevo and Rosen (2011) to bound the estimate from both sides.<sup>8</sup>

Define  $\lambda$  as the ratio between the correlation between the IV and the error term (de-

---

<sup>7</sup>Assumptions 1 and 2 are linear specification and exogenous variables for other controls.

<sup>8</sup>Precisely, the required condition for bounds in both sides are  $(\sigma_{z\tilde{x}}\sigma_x - \sigma_{x\tilde{x}}\sigma_z)\sigma_{z\tilde{x}} > 0$ , where  $z$  is the IV,  $x$  is the endogenous variable,  $\tilde{x}$  is the residual of the exogenous variable after regressing for all the other controls,  $\sigma_{z\tilde{x}}$  is the covariance between  $z$  and  $\tilde{x}$ ,  $\sigma_x$  and  $\sigma_z$  are the standard deviations of  $x$  and  $z$ , respectively. Since both  $\sigma_x$ ,  $\sigma_z$ , and  $\sigma_{x\tilde{x}}$  are positive, we only need to check out whether  $\sigma_{z\tilde{x}}$  is negative, i.e., whether the IV, conditioning on other regressors, is negatively correlated with the endogenous variable, which is confirmed in Panel B of Table 3.

noted  $\rho_{z\varepsilon}$ ) and the the correlation between the endogenous variable and the error term (denoted  $\rho_{x\varepsilon}$ ), i.e.,  $\lambda \equiv \frac{\rho_{z\varepsilon}}{\rho_{x\varepsilon}}$ . Fix a  $\lambda$ , we can construct a “perfect” IV, denoted  $v(\lambda)$ , that is correlated with the endogenous variable but not the error term using the following formula:

$$v(\lambda) = \sigma_x z - \lambda \sigma_z x, \quad (6)$$

where  $x$  and  $z$  are the endogenous variable and the IV, respectively, and  $\sigma_x$  and  $\sigma_z$  are their respective standard deviations.<sup>9</sup> The problem, however, is that we do not know the value of  $\lambda$ . Assumption 4, that the endogenous variable is at least as endogenous as the IV, implies that  $\lambda \in [0, 1]$ . In Proposition 3, Nevo and Rosen (2011) show that performing the same IV estimation using the original IV and the constructed IV with  $\lambda = 1$  (the worst case in which the IV is as endogenous as the endogenous variable), the two resulting estimated coefficients bound the value of  $\beta$ .

Table 4 shows that using the original IV, *Import penetration*<sub>*i,-c*</sub>, the estimate is 0.567, whereas using the constructed IV taking  $\lambda = 1$ , the estimate is 0.208 and is statistically significant. In other words,  $\beta$  is bounded between 0.208, and 0.567. This suggests that even if we entertain the doubt that the IV fails the orthogonal condition, our result that imports spur incremental innovation remains robust.

## 4.2 Robustness

**Technologically capable firms locating together?** An alternative explanation for our result is that technologically capable firms within an industry self-select to locate at major markets with higher import penetration. Being technologically capable, they also innovate more.

Table 5 presents conflicting evidence on whether technologically capable firms locate at where import penetration is high. Taking the industry-city level average (subscript

---

<sup>9</sup>The proof of this claim is not reported but is available upon request.

$i,c$ ), *Import penetration* does correlate positively with whether any patent is granted in the past three years (*Patent*) and whether any engineer or technician is hired in 2002 (*Engineers & technicians*), but correlates negatively with that of whether any R&D personnel is hired over the past three years (*R&D personnels*).<sup>10</sup> Table 6 directly controls for these three industry-city level measures. Given that *Import penetration* remains statistically significantly positive, our result is unlikely driven by this concern.

**Alternative measures.** Table 7 shows our results are robust to alternative measures. Column 1 uses the dummy *R&D* that takes a value of 1 if the R&D expenditure in 2002 is positive, and 0 otherwise.<sup>11</sup> Column 2 uses the percentage of overseas suppliers in terms of total competitors' output (denoted by *Importers*). Column 3 uses the industry-city level average of *Import penetration* $_{i,c}$ , which is better at averaging out any errors if firms misreports their own import penetration ratio.

**Different types of innovation.** We also investigate the association between the import penetration ratio and each of the five types of innovation. If a firm has introduced any new products/services, production processes, business lines, management practices, and quality controls over the past three years, the dummies *Product*, *Process*, *Business*, *Management* and *Quality*, respectively, take the value of 1, and 0 otherwise. However, an implicit challenge exists. While a firm can engage in different types of incremental innovation, regressing one of the above dummies with *Import penetration* may involve comparing firms having engaged in that type of innovation with non-innovators, as well as firms having engaged in that type of innovation with those having engaged in other types.

We use two approaches to address this issue. First, we run seemingly unrelated regressions to jointly estimate each of the five types of innovation. Second, we separately estimate each of the five types but drop those firms that have engaged in other types of

---

<sup>10</sup>Controlling these three firm-specific variables directly in the estimation of (1) neither diminishes the significance nor dramatically changes the size of the estimated coefficient of *Import penetration*.

<sup>11</sup>The correlation between *R&D* and *Innovation* is 0.309 and is statistically significant at the 1% level. We have also checked that *Innovation* is statistically significantly correlated with the actual R&D expenditure for the firms in the past years. These expenditure levels are available for manufacturing firms only.

innovation. Doing so ensures that innovators of one particular type of innovation are compared with non-innovators, and the cost is reduced sample size. Panels A and B of Table 8 show the first and second approach, respectively. Consistent with the above concern, the estimates are larger under the second approach. While the second approach shows a positive association between the import penetration ratio and four of the five types of innovations, the statistical significance is weaker under the first approach.

**Sample bias and attrition.** Although some firms have multiple lines of businesses, *Import penetration* only pertains to a firm's major business line in its major market. However, *Innovation* does not distinguish innovations of the major business line from those of other business lines. Are our results driven by the inclusion of firms with multiple lines of businesses? Column 1 of Table 9 contains a sub-sample of firms whose main business contributes over 90% to their total sales. The results suggest that this concern is unlikely to drive our results.

Can attrition bias drive our results? After all, one third of the firms did not report their import penetration ratios. Column 2 follows Duflo (2001) in including the polynomial terms of the propensity in the estimation, where the propensity is the estimated propensity of a firm in answering the survey question on its import penetration ratio based on firm and CEO characteristics. Column 3 uses the inverse-probability-weighted method (Wooldridge, 2002, 2007), in which the inverse of the estimated propensity is used as the sample weight in the OLS estimation. The results suggest that our results are unlikely to be driven by sample attrition bias.

## 5 The underlying mechanisms

### 5.1 Competition effects

Firms may strive to improve their production because increased foreign competition forces them to do so. This competitive effect has been documented for firms in the North (Bloom,

Draca, and Van Reenen, 2011). Do firms in the South react to foreign imports by innovating *because* of increased foreign competitive pressure? In Aghion et al. (2005), this underlying mechanism implies that increased import competition reduces pre-innovation rents more than post-innovation rents. The competitive effect is weaker if foreign competition cannot easily affect pre-innovation rents.

We exploit this idea by looking at firms' share of output sold to government agencies (denoted by *State sales*). In China, trade relationships with governmental agencies are largely determined by political connection rather than economic considerations. These sticky relationships suggest that the pre-innovation rents of firms with more sales to government agencies should be less affected by increased import competition. Accordingly, we use the estimation strategy in Rajan and Zingales (1998).

Table 10 shows that the interaction term between *Import penetration* and *State sales* has a negative and statistically significant estimated coefficient, implying that firms with more sales to government agencies are less likely to innovate as a response to imports.

A concern is whether the types of firms engaged in governmental sales are inherently different from those in market sales. We thus verify whether firm-level technological capability is strongly correlated with the measure of governmental sales. Table 11 shows no strong evidence.<sup>12</sup>

## 5.2 High-tech versus mature manufacturing industries

Table 12 follows Henderson, Kuncoro and Turner (1995) in dividing manufacturing industries into two groups: mature industries (including garment and leather products, electronic parts making, household electronics, food processing, and metallurgical products), and new high-tech industries (including electronic equipment, auto and auto parts,

---

<sup>12</sup>We also regress *State sales* with each of the three technological capability measures, with and without controlling for all the other controls. The results suggest no unified relationship.

chemical products and medicine, and biotech products and Chinese medicine).<sup>13</sup> The discouraging effect of Aghion et al. (2005) dominates if the laggard firms are further away from the technological frontier. We find a consistent pattern that *Import penetration* is positive and significant only for firms in mature manufacturing industries where the technology gap with developed countries is narrow. However, in high-tech industries where technological gap is huge, we find insignificant and even negative estimated coefficient.

## 6 Conclusion

The South has increasingly engaged in innovations, which have significant implications on both the global science landscape and the international trade pattern. This paper documents that imports are robust determinants of their incremental innovations. This finding echoes with that on firms in the North. Specifically, the firm survey by the World Bank in China shows that firms facing higher import penetration are more likely to engage in incremental innovations.

Our paper focuses on ruling out other alternative explanations. Using different strategies, along with a battery of robustness checks, our paper documents that the result is unlikely to be attributed to other global connections, endogeneity, or the clustering of technologically capable firms in markets with intense import competition. We also find that firms in the South do respond to imports by innovating because foreign competition forces them to improve their production, an underlying mechanism for understanding the responses of firms in the North.

---

<sup>13</sup>As the industrial classifications in China are different from those in the US, an effort is made to match the five mature industries and four high-tech industries used by Henderson et al. (1995) with the corresponding industries in China.

## References

- [1] **Aghion, Philippe, Nicholas Bloom, Richard Blundell, Rachel Griffith, and Peter Howitt.** 2005. "Competition and Innovation: An Inverted-U Relationship." *Quarterly Journal of Economics*, 120(2): 701-728.
- [2] **Ayyagari, Meghana, Asli Demirguc-Kunt, and Vojislav Maksimovic** 2008. "Formal versus informal finance: Evidence from China." *Policy Research Working Paper Series 4465, The World Bank.*
- [3] **Bloom, Nicholas, Mirko Draca, and John Van Reenen.** 2011. "Trade Induced Technical Change? The Impact of Chinese Imports on Innovation, Diffusion and Productivity." *NBER Working paper no. 16717.*
- [4] **Bloom, Nicholas, Paul Romer, and John Van Reenen.** 2010. "A Trapped Factor Model of Innovation." *Working paper.*
- [5] **Cull, Robert, and Lixin Colin Xu.** 2005. "Institutions, Ownership, and Finance: the Determinant of Profit Reinvestment among Chinese Firms." *Journal of Financial Economics*, 77(1): 117-146.
- [6] **Duflo, Esther.** 2001. "Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment." *American Economic Review*, 91(4): 795-813.
- [7] **Gorodnichenko, Yuriy, Jan Svejnar, and Katherine Terrell.** 2010. "Globalization and Innovation in Emerging Markets." *American Economic Journal: Macroeconomics*, 2(2): 194-226.
- [8] **Henderson, Vernon, Ari Kuncoro, and Mathew Turner.** 1995. "Industrial Development in Cities." *Journal of Political Economy*, 103(5): 1067-90.

- [9] **Hausman, Jerry, Gregory Leonard and J. Douglas Zona.** 1994. "Competitive Analysis with Differentiated Products." *Annales D'Economie Et De Statistique*, 34: 159-180.
- [10] **Leamer, Edward E.** 2007. "A Flat World, A Level Playing Field, a Small World After All, or None of the Above?" Review of *The World is Flat*, by Thomas J. Friedman, *Journal of Economic Literature*, March (2007).
- [11] **Li, Hongbin, Lingsheng Meng, Qian Wang and Li-an Zhou.** 2008. "Political Connections, Financing and Firm Performance: Evidence from Chinese Private Entrepreneurs." *Journal of Development Economics*, 87(2): 283-299.
- [12] **Lin, Chen, Ping Lin, and Frank Song.** 2010. "Property Rights Protection and Corporate R&D: Evidence from China." *Journal of Development Economics*, 93(1): 49-62.
- [13] **Merlitz, Marc.** 2011. "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity." *Econometrica*, 71(6): 1695-1725.
- [14] **Nevo, Aviv, and Adam Rosen.** 2011. "Identification with Imperfect Instruments." *Review of Economics and Statistics*, forthcoming.
- [15] **OECD.** 2008. *Science, Technology and Industry, 2008*. Paris: Organisation for Economic Co-operation and Development Press.
- [16] **OECD.** 2006. "South-South Trade: Vital for Development." *Policy Brief*, August 2006. Paris: Organisation for Economic Co-operation and Development Press.
- [17] **Rajan, Raghuram G., and Luigi Zingales.** 1998 "Financial Dependence and Growth." *American Economic Review*, 88(3): 559-86.
- [18] **Rodrik, Dani.** 2006. "What's So Special About China's Exports?" *China and World Economy*, 14(5): 1-19.
- [19] **Schott, Peter K.** 2008. "The Relative Sophistication of Chinese Exports." *Economic Policy* 23(53):5-49.

- [20] **Staiger, Douglas, and James H. Stock.** 1997. "Instrumental Variables Regression with Weak Instruments." *Econometrica*, 65(3): 557-586.
- [21] **Teshima, Kensuke.** 2009. "Import Competition and Innovation at the Plant Level: Evidence from Mexico." *Working Papers*.
- [22] **UNCTAD.** 2005. *Trade and Development Report 2005*. United Nations publication, New York and Geneva.
- [23] **Wooldridge, Jeffrey M.** 2002. "Inverse Probability Weighted M-estimators for Sample Selection, Attrition, and Stratification." *Portuguese Economic Journal*, 1(2): 117-139.
- [24] **Wooldridge, Jeffrey M.** 2007. "Inverse Probability Weighted M-Estimation for General Missing Data Problems." *Journal of Econometrics*, 141(2): 1281-1301.
- [25] **World Trade Organization.** 2005. "Developing countries' goods trade share surges to 50-year peak," April 14, Press Release/401.

## **A Data appendix: the Survey of Chinese Enterprises (SCE)**

The SCE was conducted in 2003 by the World Bank. Cull and Xu (2005) and Lin et al. (2010) also use the SCE. The newest wave of SCE was conducted in 2005, covering Chinese firms in 100+ cities. However, this wave does not include the key variables used in this paper.

The SCE randomly-surveyed 2,400 firms across 18 cities and 14 industries. For a balanced representation, the SCE selects 18 cities from 16 provinces located in five areas in China: Northeast area – Benxi, Changchun, Dalian, and Harbin; Coastal area – Hangzhou, Jiangmen, Shenzhen, and Wenzhou; Central area – Changsha, Nanchang, Wuhan, and Zhengzhou; Southwest area – Chongqing, Guiyang, Kunming, and Nanning; and Northwest area – Lanzhou and Xi’an. In each city, the SCE randomly samples 100 or 150 firms from 14 industries, including 9 manufacturing (garment and leather products, electronic equipment, electronic parts making, household electronics, automobile and automobile parts, food processing, chemical products and medicine, biotech products and Chinese medicine, and metallurgical products) and 5 service industries (information technology, accounting and non-banking financial services, advertisement and marketing, business services, and transportation equipment including telecommunication and ship building).

The SCE is composed of two parts. One is a general questionnaire directed to the senior management seeking information about the firm, innovation, product certification, marketing, relation with suppliers and customers, access to markets and technology, relation with the government, labor, infrastructure, international trade, finance, taxation, and the CEO and board of directors. The other questionnaire is directed to the accountant and personnel manager covering ownership, various financial measures, and labor and training. Most of the information from the first part of the SCE pertains to 2002, whereas the second part pertains to the period of 2000-2002.

Table 1: Summary statistics

Variable	Obs	Mean	Std. Dev.	Min	Max	Description
Innovation	2357	0.693	0.461	0	1	= 1 if introduced either new products/services, businessline, process, mgmt. techniques, or quality control
Import penetration	1559	0.054	0.120	0	0.693	$\ln(1 + \text{imports as a \% of total sales at the firm's major market})$
<b>Firm characteristics</b>						
<i>Conventional set</i>						
Labor productivity	2351	4.154	1.552	-4.254	16.325	$\ln(\text{total sales} / \text{total employment})$ in 2000
Firm size	2383	4.816	1.542	0	11.358	$\ln(\text{total employment})$ in 2000
Firm age	2400	2.430	0.799	1.099	3.970	$\ln(\text{years of establishment up to the end of 2002})$
Capital stock	2290	8.536	2.522	0	17.427	$\ln(\text{total fixed assets in book value})$ in 2000
State ownership	2399	0.241	0.417	0	1	% of shares owned by the government in 1999
Degree of computerization	2367	33.330	34.559	0	100	% of workforce regularly uses a computer in their jobs
<i>Institutional qualities</i>						
Property rights protection	2225	0.343	0.314	0	1	% of government officials that is oriented toward helping rather than hindering firm operations
Contract enforcement	2068	0.640	0.389	0	1	% that the legal system will uphold the firm's contractual and property rights in business disputes
<i>Foreign connections</i>						
Foreign share	2399	0.059	0.168	0	0.693	$\ln(1 + \%$ shares owned by foreign individuals, institutional investors, firms, and banks in 1999)
Overseas customers	2326	0.091	0.260	0	1	% of oversea customers (measured by sales)
Overseas suppliers	2212	0.053	0.177	0	1	% of oversea suppliers (measured by expenditures)
<b>CEO characteristics</b>						
<i>Human capital</i>						
CEO education	2382	15.643	2.394	0	19	CEO's years of schooling
CEO tenure	2371	5.771	4.255	1	33	Years of being CEO
Deputy CEO previously	2378	0.274	0.446	0	1	= 1 if the CEO was the deputy CEO before becoming the CEO
<i>Political capital</i>						
Government cadre previously	2378	0.060	0.237	0	1	= 1 if the CEO was the government cadre before becoming the CEO
Party member	2351	0.668	0.471	0	1	= 1 if the CEO was the member of the Chinese Communist Party
<b>IV</b>						
Import penetration <sub><i>i, -c</i></sub>	2400	0.052	0.027	0.005	0.126	Average of <i>Import penetration</i> within the same industry across different cities
<b>Alternative measures</b>						
R&D	1573	0.418	0.493	0	1	= 1 if R&D expenditure in 2002 $\geq 0$ , = 0 otherwise
Importers	2274	0.560	1.239	0.000	4.615	$\ln(1 + \%$ of overseas competitors (in terms of output))
<b>Others</b>						
Agglomeration	2274	0.281	0.274	0	0.693	% of output in the firm's major market supplied by competitors within the same city
State sales	2206	0.046	0.144	0	1	% of sales sold to governmental agencies
Patent	2354	0.102	0.303	0	1	= 1 if the firm has been granted a patent in the past 3 years
Engineers & technicians	2351	0.775	0.417	0	1	= 1 if the firm has hired at least one engineer or technician in 2002
R&D personnels	2400	0.630	0.483	0	1	= 1 if the firm has hired at least one R&D staff in the past 3 years

Table 2: Main results

	1	2	3	4	5
Panel A: OLS Estimation	Dependent Variable: Innovation				
Import penetration	0.355*** [0.086]	0.279*** [0.087]	0.181** [0.085]	0.191** [0.087]	0.190** [0.085]
<b>Firm characteristics</b>					
<i>Global connections</i>					
Foreign share			-0.006 [0.068]	0.028 [0.062]	0.024 [0.065]
Overseas customers			0.058 [0.045]	0.062 [0.051]	0.075 [0.055]
Overseas suppliers			-0.009 [0.068]	-0.04 [0.066]	-0.041 [0.065]
<i>Conventional set</i>					
Labor productivity			0.027*** [0.010]	0.028*** [0.010]	0.026*** [0.010]
Firm size			0.044*** [0.013]	0.039*** [0.013]	0.033** [0.013]
Firm age			-0.082*** [0.018]	-0.087*** [0.019]	-0.082*** [0.019]
Capital stock			0.005 [0.010]	0.004 [0.010]	0.002 [0.010]
State ownership			0.056* [0.031]	0.05 [0.032]	0.046 [0.032]
Degree of computerization			0.002*** [0.000]	0.001** [0.000]	0.001** [0.000]
Agglomeration					-0.095* [0.051]
<i>Institutional qualities</i>					
Property rights protection			-0.032 [0.037]	-0.04 [0.038]	-0.025 [0.036]
Contract enforcement			0.028 [0.030]	0.039 [0.029]	0.032 [0.030]
<b>CEO characteristics</b>					
<i>Human capital</i>					
CEO education				0.019*** [0.007]	0.017** [0.007]
CEO tenure				0.003 [0.003]	0.003 [0.003]
Deputy CEO previously				0.050** [0.022]	0.050** [0.023]
<i>Political capital</i>					
Government cadre previously				-0.04 [0.062]	-0.054 [0.062]
Party member				-0.024 [0.026]	-0.019 [0.026]
<b>Dummies</b>					
Competition dummies	No	No	No	No	Yes
Industry Dummies	No	Yes	Yes	Yes	Yes
City Dummies	No	Yes	Yes	Yes	Yes
Number of Observations	1538	1538	1291	1250	1230
R-squared	0.0094	0.1046	0.1552	0.1704	0.175
F-test	17.06	8.714	8.132	10.76	10.83
p-value for F-test	0.0000	0.0000	0.0000	0.0000	0.0000
Panel B: Probit Estimation	Dependent Variable: Innovation				
Import penetration	0.402*** [0.116]	0.341*** [0.115]	0.232** [0.113]	0.256** [0.119]	0.259** [0.118]

Note: White-robust standard errors clustered at the industry-city level are reported in brackets. \*, \*\* and \*\*\* represent statistical significance at the 10%, 5%, and 1% levels, respectively. A constant is included in each regression. Panel B reports the marginal effect of Probit estimation.

Table 3: Instrumental variable estimation

	1	2
Panel A: 2nd stage of IV estimation	Dependent Variable: Innovation	
Import penetration	0.802** [0.356]	0.567** [0.271]
<b>Firm characteristics</b>	No	Yes
<b>CEO characteristics</b>	No	Yes
<b>Dummies</b>		
Competition dummies	No	Yes
Industry Dummies	Yes	Yes
City Dummies	Yes	Yes
Panel B: 1st stage of IV estimation	Dependent Variable: Import penetration	
Import penetration <sub><i>i, -c</i></sub>	-4.667*** [0.645]	-5.822*** [0.735]
Anderson canonical correlation LR statistics	22.305	20.057
Cragg-Donald F-statistics	52.428	62.667
Number of Observations	1538	1230
Panel C: IV-Probit	Dependent Variable: Innovation	
Import penetration	0.932*** [0.343]	0.799*** [0.314]

Note: White-robust standard errors clustered at the industry-city level are reported in brackets. \*\* and \*\*\* represent statistical significance at the 5% and 1% levels, respectively. A constant term is included in each regression. The first stages of the IV estimations include the same set of control variables as in the corresponding second stage, but the estimated coefficients of these control variables are not reported to save space. Panel C reports the marginal effect of IV-Probit estimation.

Table 4: Bounding parameters: relaxing exclusion restriction

	1	2
Panel A: 2nd stage of IV estimation	Dependent Variable: Innovation	
IV	Original IV	Constructed IV
Import penetration	0.567** [0.271]	0.208** [0.083]
<b>Firm characteristics</b>	Yes	Yes
<b>CEO characteristics</b>	Yes	Yes
<b>Dummies</b>		
Competition dummies	Yes	Yes
Industry Dummies	Yes	Yes
City Dummies	Yes	Yes
Number of Observations	1230	1230

White-robust standard errors clustered at the industry-city level are reported in brackets. \*\* represents statistical significance at the 5% level. A constant is included in each regression. The first stages of the IV estimations include the same set of control variables as in the corresponding second stage.

Table 5: Pairwise correlations at the industry-city level

	Import penetration $_{i,c}$	Patent $_{i,c}$	Engineering & technicians $_{i,c}$	R&D personnels $_{i,c}$
Import penetration $_{i,c}$	1			
Patent $_{i,c}$	0.1618***	1		
Engineering & technicians $_{i,c}$	0.2524***	0.5289***	1	
R&D personnels $_{i,c}$	-0.1088***	-0.1462***	-0.3037***	1

\* \* \* represents statistical significance at the 1% level. Subscript  $i, c$  imply the average of all firms within an industry-city.

Table 6: Estimation with controls on the industry-city level technological capability

	1	2	3
Dependent Variable: Innovation			
Import penetration	0.536** [0.272]	0.439* [0.265]	0.577** [0.279]
Patent $_{i,c}$	-0.135 [0.146]		
Engineering & technicians $_{i,c}$		0.264** [0.112]	
R&D personnels $_{i,c}$			0.155** [0.072]
<b>Firm characteristics</b>	Yes	Yes	Yes
<b>CEO characteristics</b>	Yes	Yes	Yes
<b>Dummies</b>			
Competition dummies	Yes	Yes	Yes
Industry Dummies	Yes	Yes	Yes
City Dummies	Yes	Yes	Yes
Number of Observations	1230	1230	1230

White-robust standard errors clustered at the industry-city level are reported in brackets. \* and \*\* represent statistical significance at the 10% and 5% levels, respectively. A constant is included in each regression.

Table 7: Robustness check: Alternative measures

	1	2	3
Dependent variable	R&D	Innovation	Innovation
Import penetration	0.240* [0.136]		
Importers		0.188** [0.078]	
Import penetration <sub><i>i,c</i></sub>			0.400*** [0.137]
<b>Firm characteristics</b>	Yes	Yes	Yes
<b>CEO characteristics</b>	Yes	Yes	Yes
<b>Dummies</b>			
Competition dummies	Yes	Yes	Yes
Industry Dummies	Yes	Yes	Yes
City Dummies	Yes	Yes	Yes
Number of Observations	915	1654	1618

Note: White-robust standard errors clustered at the industry-city level are reported in brackets. \*, \*\* and \*\*\* represent statistical significance at the 10%, 5%, and 1% levels, respectively. A constant is included in each regression. Subscript  $i, c$  imply *Import penetration* is calculate as its average of all firms within an industry-city unit.

Table 8: Breaking down the 5 types of innovations

	1	2	3	4	5
Dependent variable	Product	Process	Business	Management	Quality
Panel A: Seemingly unrelated regressions					
Import penetration	0.163 [0.112]	0.196* [0.111]	0.0421 [0.104]	0.155 [0.105]	0.0582 [0.123]
No. of observation	1239	1239	1239	1239	1239
Panel B: Innovators versus non-innovators					
Import penetration	0.173 [0.108]	0.232* [0.118]	0.272* [0.147]	0.196** [0.097]	0.230** [0.104]
No. of observation	879	781	636	1042	1015
<b>Controls</b>					
<b>Firm characteristics</b>	Yes	Yes	Yes	Yes	Yes
<b>CEO characteristics</b>	Yes	Yes	Yes	Yes	Yes
<b>Dummies</b>					
Competition dummies	Yes	Yes	Yes	Yes	Yes
Industry dummies	Yes	Yes	Yes	Yes	Yes
City dummies	Yes	Yes	Yes	Yes	Yes

Note: White-robust standard errors clustered at the industry-city level are reported in brackets. \* and \*\* represent statistical significance at the 10% and 5% levels, respectively. A constant is included in each regression. Panel A reports seemingly unrelated regression results, while Panel B reports OLS results.

Table 9: Robustness check: Sample bias and sample attrition

Sample	1 Firms with focused business	2 All firms	3 All firms
Dependent variable: Innovation			
Import penetration	0.153* [0.091]	0.193** [0.084]	0.186** [0.090]
Probability (reporting import penetration)		0 [0.000]	
Probability (reporting import penetration) <sup>2</sup>		-0.034* [0.018]	
<b>Firm characteristics</b>	Yes	Yes	Yes
<b>CEO characteristics</b>	Yes	Yes	Yes
<b>Dummies</b>			
Competition dummies	Yes	Yes	Yes
Industry Dummies	Yes	Yes	Yes
City Dummies	Yes	Yes	Yes
Number of Observations	755	1230	1230

Note: White-robust standard errors clustered at the industry-city level are reported in brackets. \* and \*\* represent statistical significance at the 10% and 5% levels, respectively. A constant is included in each regression. Column 1 includes only firms with their major line of businesses accounting for more than 90% of their total sales. Column 2 includes the polynomial terms, while Column 3 uses the inverse of the estimated propensity of reporting import penetration ratio as sample weight.

Table 10: Heterogeneous response estimation

Dependent Variable: Innovation	
Import penetration	0.284*** [0.078]
State sales	0.328*** [0.069]
Import penetration x State sales	-2.154** [0.939]
<b>Firm characteristics</b>	Yes
<b>CEO characteristics</b>	Yes
<b>Dummies</b>	
Competition dummies	Yes
Industry Dummies	Yes
City Dummies	Yes
Number of Observations	1216
R-squared	0.1849
F-test	11.75
p-value for F-test	0

Note: White-robust standard errors clustered at the industry-city level are reported in brackets. \*\* and \*\*\* represent statistical significance at the 5% and 1% levels, respectively. A constant is included in the regression.

Table 11: Pairwise correlations at the firm level

	State sales	Patent	Engineers & technicians	R&D personnels
State sales	1			
Patent	0.0009	1		
Engineers & technicians	-0.0082	0.1423***	1	
R&D personnels	0.0787***	0.0867***	-0.0271	1

\*\*\* represents statistical significance at the 1% level.

Table 12: When do imports matter?

Sample	1 Mature industries	2 High-tech industries
Dependent variable: Innovation		
Import penetration	0.267*** [0.100]	-0.044 [0.152]
<b>Firm characteristics</b>	Yes	Yes
<b>CEO characteristics</b>	Yes	Yes
<b>Dummies</b>		
Competition dummies	Yes	Yes
Industry Dummies	Yes	Yes
City Dummies	Yes	Yes
Number of Observations	525	388
R-squared	0.2088	0.19

Note: White-robust standard errors clustered at the industry-city level are reported in brackets. \*\*\* represents statistical significance at the 1% level. A constant is included in each regression.