

Do Imports Spur Incremental Innovation in the South?

Yi Lu* and Travis Ng^{†‡}

June 2011

Abstract

We estimate 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. Following Aghion, Bloom, Blundell, Griffith and Howitt (2005), 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: F23, M16, O33.

*The National University of Singapore. Email: ecluyi@nus.edu.sg

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

‡We would like to thank Jiahua Che, Zhigang Tao, and Tat-kei Lai for their helpful comments.

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 product/services, business line, management practices, quality control, or production process in the past three years.¹ Different from the traditional import penetration ratio measured at the industry-level, our firm-level 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 been much more globalized too. Linking their globalization trend with their innovation may draw potentially important policy implications.² Third, Leamer (2007) argues that the unintended transfer of innovative processes in manufacturing alongside outsourcing of production tops all other concerns of “the world is flat”; it deserves the developed world’s attention to examine their causes and consequences.³ 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 too. While the developed world may have recognized their outsourcing brings

¹“Process innovations” in the literature refers to small-step cost-cutting innovations. In contrast, “incremental innovations” embodies more than just innovations on processes; it includes product innovations, new quality control, new management practices, etc. Williamson and Zeng (2007) give examples of Chinese firms on such innovations, e.g., tailoring existing technology to niche markets at a large volume. Haier, for instance, gains a 60% share of the world’s wine-storage fridges, a niche market in fridge manufacturing. Lin and Saggi (2002) give a theory behind the interplay between process innovation, product innovation, and market competition.

²For example, increasingly vocal in opposing imports is not uncommon both in the developed countries, as well as in the developing countries. Recognizing imports as an innovation driver can be important.

³Leamer (2007)’s concern is framed in the US context: given that manufacturing, farming, and transportation only accounts for 15% of the current GDP of US and that innovative processes are attached to production, he asks how much is the latest financial derivative really worth and how much is it going to change the lives of the Americans?

innovation over to the developing world, it is interesting to ask instead whether the fact that they also are exporting more to the developing countries also spurs their innovativeness.

The strong association between import penetration and incremental innovation can also be due to the possibility that more innovative firms are more likely to have other forms of significant global connection. Foreign ownership may be more prevalent among them, and they may be more likely to trade with the rest of the world. Even if imports do not spur incremental innovation, to the extent that these firms are more likely to locate their major markets at where foreign competition are intense, the strong association merely reflects other global connection. Our dataset allows us to rule out this alternative explanation by directly controlling for firms' foreign ownership, their fractions of sales abroad, and their fraction of imported inputs.

Another alternative explanation is that more capable firms both innovate more and are more likely to position their major markets at where foreign competition is more intense. We use four strategies to show that this alternative explanation is unlikely to be sufficient to drive our results.

Our first strategy is to directly control for firm and CEO characteristics, most notably their lagged labor productivity, their levels of capital, and the employment size. We also pay attention to the CEO's education level and his political capital.

Our second strategy is to address any omitted firm-level variables by instrumenting our regressor of interest, the firm-level import penetration ratio, with an instrumental variable (IV) that is unlikely to be correlated with any omitted firm-level variables. Our IV is the average of import penetration ratios faced by firms belonging to the same industry but located in other cities. Section 4.1 details the relevance of this IV and why it is unlikely to fail the exclusion restriction. Section 4.1.1 shows a falsification test that further lends confidence on the IV.

Our third strategy is to use the imperfect IV inference method developed in Nevo and Rosen (2011). The idea is to relax the IV's identifying assumption by assuming it to be

endogenous too, but less so than our regressor of interest. Indeed, despite relaxing the identifying assumption, the significant association between import penetration ratio and incremental innovation remains robust.

Our fourth strategy is to address the claim that technologically capable firms may cluster at major markets that face more intense import competition. We fail to find compelling evidence supporting this claim.

The four strategies, along with a battery of other robustness checks using alternative measures and alternative specification, and controlling for sample attrition bias, provide compelling evidence that the alternative explanations and other econometric errors 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 better their goods/services for higher rent) is an important channel through which import penetration ratio affects firms' incremental innovation. If competition effect is a significant channel, those firms with pre-innovation rents less influenced by foreign imports should have a weaker correlation between their innovation and imports. We exploit this idea with three strategies. First, Section 5.1 uses the heterogenous response estimation à la Rajan and Zingales (1998) to show that such an empirical pattern indeed exists. Those firms selling more to government agencies should have their pre-innovation rents less influenced by foreign imports. Their innovation are also significantly less correlated with imports.

Our second strategy is to contrast dominant and non-dominant firms. The intuition is that compared to non-dominant firms, dominant firms with a bigger market share are relatively more likely to possess certain inherent attributes that sustain rents. Their pre-innovation rents are thus less influenced by imports. We find that dominant firms' innovation correlates less with imports than non-dominant firms.

Our third strategy is to contrast manufacturing and services industries, basing on the intuition that services are less trade-able than manufacturing outputs.

These three strategies provide indirect evidence suggesting that, like in developed countries, competition effect is likely to be an important channel through which import penetration ratio affects firms' incremental innovation.

We also show evidence consistent with the discouragement effect (domestic firms give up competing with foreign imports for the imports are way better than theirs) in Aghion, Bloom, Blundell, Griffith and Howitt (2005). Firms in the mature manufacturing industries that have their products relatively closer to their respective global technological frontiers respond to imports to innovate, but not those firms in the high-tech manufacturing industries that have their products relatively far away their respective global technological frontiers.

Our paper echoes well with Gorodnichenko, Svejnar and Terrell (2010). They document that firms in European developing countries also respond to imports by engaging in more small-step innovation. We strengthen their results first by estimating too a significant effect with the size in the neighborhood of their estimates, and second, by extending their results from emerging European economies to China. China is an important context not to be missed in examining the link between imports and innovation in the South. It tops among other developing countries in terms of its economic significance of innovation.⁴ It spends the most in R&D among them.⁵ While a world exporter, China is also a gigantic importer.⁶

Teshima (2009) measures import competition faced by Mexican firms by firm-specific import tariff reduction. He shows that tariff reduction raises their spending on both process R&D and product R&D. While his main focus is R&D spending, ours is the innovation output rather than the input. While we use import competition directly reported, he creatively derives firm-specific tariff reductions from firms' output data. To the extent that more R&D

⁴Between 1990-1992 and 2000-2002 periods, China's share of innovative new goods in US import increased from 17% to 48%, or from 0% to 5% in terms of trade value.

⁵China's expenditure on R&D reached USD 86.8 billion in 2006 after an impressive growth at around 19% annually in real terms from 2001 to 2006 (OECD, 2008). It claimed 2% of the global share of total R&D expenditure in 1996, but dramatically rose to 7.5% in 2005, just behind the US (~ 35%), the EU27 (~ 24%), and Japan (~ 14%).

⁶The 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).

spending produces more innovation stochastically, our result is closely in-line with his.

2 Linking imports with incremental innovation in the South

Aghion, Bloom, Blundell, Griffith and Howitt (2005) show competition can either promote or discourage innovation. If increased competition reduces pre-innovation rents more than it reduces 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 and thus firms are discouraged from innovating (the discouraging effect). The competitive effect likely dominates if a firm operates at a technological level on par with 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 laggard relative to their import competitors. They are *not* at technological par with their import competitors because North-North trade and North-South trade dominates South-South trade (OECD, 2006; UNCTAD, 2005).⁷ Thus the discouraging effect may dominate in the South.

Their framework shows that 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, competition effects dominates, and vice versa. While the South do import from the rest of the world, their trade barriers are considerably higher than those in the North.⁸ 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

⁷The OECD (2006)’s estimates merchandise exports and trade in services among the South in 2002 is around 6% and 10%, respectively, of the corresponding world figures. UNCTAD (2005) gives much in-depth analysis of the composition and the trend of South-South trade.

⁸OECD (2006) estimates a simple average of tariff rates in 2001 for the North importers are 4.4 and 5.0 for exports from the North and the South, respectively, relative to a higher average of 9.9 and 11.1 for the South importers.

dominate instead.

While Aghion, Bloom, Blundell, Griffith and Howitt (2005) proxy innovation with patenting, Gorodnichenko, Svejnar and Terrell (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, there are methods other than patenting that protect intellectual properties. Since the South has relatively weaker formal institutions, firms there are more likely to use these other methods. Besides, incremental innovation involves small-step innovation, which are less likely to justify the legal fee of patenting.

Incremental innovation of the South can thus be relevant to the exogenous parameter h (the rate of “laggard firms” to move one-step forward along the technology line by copying others’ technology) in the model of Aghion, Bloom, Blundell, Griffith and Howitt (2005). If there are more imports, most of them from more advanced countries, this copying effect can be strong (increased imports may raise h).⁹ The relatively more prevalent North-South trade than South-South trade implies the copying effect can be significant in the South. In contrast to Bloom, Romer and Van Reenen (2010) where firms in the North innovate in face of imports from the South because the opportunity cost of producing new goods falls (imports are old goods), firms in the South may innovate in face of imports because copying directly lowers the cost of doing so.

How increased imports affect incremental innovation in the South is ultimately an empirical question. The competitive and the copying effects suggest a positive relationship, while the discouraging effect a negative one. The nature of incremental innovation and the trade pattern of the South reflect, interestingly, some relevance of each of these effects in ways rather different from firm and industry in the North.

⁹A suspicious empirical pattern consistent with the copying effect is that China’s exports to the US overlap to a surprising extent with those from the OECDs (Schott, 2008) and China’s exports are sophisticated as those of a country three times richer (Rodrik, 2009).

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. Gorodnichenko, Svejnar and Terrell (2010) also use cross-sectional analysis on firm-level survey data conducted by the World Bank. This section introduces the variables and summarizes their strengths and shortcomings, and describes how these shortcomings bias our estimation.

3.1 Incremental innovation

The SCE captures the concept of incremental innovations closely by asking whether the firm has introduced any new product/services, new business line, new management practice, new quality control, and new production process in the past three years. The dummy *Innovation* is 1 if the firm answers yes in any of these questions, and is 0 if otherwise. Table 1 shows slightly less than 70% of the firms have engaged in either one of the 5 types of incremental innovations.

Our measure *Innovation* cannot be entirely free from measurement errors. First, the 5 categories cover a substantial portion, but not all possible innovation a firm could take. Second, so long as different firms perceived the term “new” differently, this self-reported variable *Innovation* is likely to be more noisy than an otherwise ideal measure of innovation that is objectively defined, and audited by researchers across firms. Conditional on our controls, if these extra noises are random, they would not bias our slope estimate but it would inflate its standard error. In turn, a larger standard error works against us by making it less likely to find 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 if otherwise. This alternative measure is available for manufacturing firms only. Around 42% of the manufacturing firms had non-zero R&D spending. Similar to the measure in Teshima (2009), this measure differs

from our measure *Innovation* by looking at the input rather than the output of research.

Table 1: Summary statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
Innovation	2357	0.693	0.461	0	1
Import penetration	1559	0.054	0.120	0	0.693
Firm characteristics					
<i>Conventional set</i>					
Labor productivity	2351	4.154	1.552	-4.254	16.325
Firm size	2383	4.816	1.542	0	11.358
Firm age	2400	2.430	0.799	1.099	3.970
Capital stock	2290	8.536	2.522	0	17.427
State ownership	2399	0.241	0.417	0	1
Degree of computerization	2367	33.330	34.559	0	100
<i>Institutional qualities</i>					
Property rights protection	2225	0.343	0.314	0	1
Contract enforcement	2068	0.640	0.389	0	1
<i>Foreign connections</i>					
Foreign share	2399	0.059	0.168	0	0.693
Overseas customers	2326	0.091	0.260	0	1
Overseas suppliers	2212	0.053	0.177	0	1
CEO characteristics					
<i>Human capital</i>					
CEO education	2382	15.643	2.394	0	19
CEO tenure	2371	5.771	4.255	1	33
Deputy CEO previously	2378	0.274	0.446	0	1
<i>Political capital</i>					
Government cadre previously	2378	0.060	0.237	0	1
Party member	2351	0.668	0.471	0	1
IV					
Import penetration _{<i>i, -c</i>}	2400	0.052	0.027	0.005	0.126
Alternative measures					
R&D	1573	0.418	0.493	0	1
Importers	2274	0.045	0.131	0	0.693
Others					
Agglomeration	2274	0.281	0.274	0	0.693
State sales	2206	0.046	0.144	0	1
Big player	2400	0.150	0.357	0	1
Medium player	2400	0.223	0.416	0	1
Patent	2354	0.102	0.303	0	1
Engineers & technicians	2351	0.775	0.417	0	1
R&D personnels	2400	0.630	0.483	0	1

3.2 Imports

We measure imports at the firm level - the variable *Import penetration* is the 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. Our direct measure of foreign competition differs from that in Gorodnichenko, Svejnar and Terrell (2010). They use firms' responses to the perceived severity of foreign competition on a 1-4 scale. Despite different measures, our estimate is shown later to be close to their estimates.

We also construct an alternative measure - *Importers* equals the logarithm 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?” Since both figures are self-reported, checking their consistency is crucial. Table 1 indicates that firms on average have 5.69% of their competitors being importers, which is in the neighborhood of the 6.43% of import penetration.

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. On the one hand, a more dis-aggregated industry-level measure of import penetration avoid lumping intermediate imports from processed imports. While processed imports elevate competitive pressure, forcing firms to innovate, imported intermediate imports actually help firms innovate in a very different ways.¹⁰ On the other hand, most firms engage in more than one industry. Classifying firms into industry can sometimes be arbitrary, more so under more dis-aggregated industry classification. Such a trade-off can be avoided by using firm-level import penetration ratios that tie imports more closely with import competition. In addition, it avoids the problem of implicitly assuming that all firms in the same industry are exposed to the same degree of foreign competition. Suppose firm A focuses on the inland regions while firm B in the same industry focuses on the coastal regions, assuming they both 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. Besides, the extent of a major market can also vary across firms; some refers to the national market, while others may be regional instead. Besides, firms with more than one major market may report *Import penetration* by arbitrarily aggregating imports across their major markets. Finally, firms

¹⁰An example was vividly discussed in Wang and Wei (2010): “Both the United States and China may export notebook computers, but Chinese manufacturers may have to import the computers most sophisticated components, such as processors (CPUs) made by Intel or ADM in the United States. In such a case, Chinese producers may specialize in the unsophisticated stage of production, although the final product is classified as sophisticated. If one were able to classify a product further into its components, China and developed countries might be found to produce different components. That is, they do not compete directly with each other. ”

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 - objectively defined firm market with researcher-audited import figures. So long as these measurement errors are random after conditioning on our controls, they would bias our slope estimate towards zero and work against us finding a significant relationship between imports and innovation.

Shifting major markets in response to imports is a 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 do take seriously the possible endogeneity by controlling for a host of variables, and by using an IV and the technique of imperfect IV. Besides, 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 the firms' global connection. As in Gorodnichenko, Svejnar and Terrell (2010), we control for *Overseas customers* (percentage of overseas customers), and *Overseas suppliers* (percentage of overseas suppliers). In addition, we control for *Foreign share* (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 engage in markets that are more likely to face higher import penetration, we directly control for a host of firm and CEO characteristics that proxy individual firms' technological capability. Firm characteristics are – *Labor productivity* (log of sales-over-

employment in 2000), *Capital stock* (log of total fixed assets in 2000), *Firm size* (log of employment), *Firm age* (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); institutional quality – *Property rights protection* (percentage of government officials oriented towards 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).¹¹

The CEO characteristics are measures of his/her human capital – *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); and his/her political capital – *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).¹²

Bloom, Romer and Van Reenen (2010) stress it is crucial to control for labor and product market for 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), and competition dummies corresponding to the question on the number of competitors of the firm’s main business line in its major market¹³

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

¹¹Lin, Lin, and Song (2010) use these variables to investigate the impacts of institutional qualities on innovation activities.

¹²Li, Meng, Wang, and Zhou (2008) use these variables to examine the impact of political connections on business performance.

¹³Ayyagari, Demircuc-Kunt, and Maksimovic (2008) use the same measures.

(denoted by *State sales*).

We also construct market share dummies. If a firm has a substantial market share (larger than 40%), the dummy *Big player* is 1, and 0 if otherwise. If a firm has a significant market share (between 10% to 40%), the dummy *Medium player* is 1, and 0 if otherwise. Table 1 shows that about 15% and 22.3% of the firms are big players and medium players, respectively.

4 Empirical Analysis

We estimate the following equation:

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

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 the results.

Columns 3-4 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.¹⁴ Column 5 controls for local competition to distinguish from foreign competition. The significantly positive association persists across all columns.

Panel B takes into account *Innovation* is a dummy by reporting the corresponding marginal effects of *Import penetration* (evaluated at means) using Probit. A similar pat-

¹⁴For the control variables, we find that larger and younger firms, and firms with higher labor productivity, are less agglomerated, with higher human capital as proxied by the average usage of computers among the workforce, and more capable CEO, innovate more. One may wonder why global connection is not significantly correlated with innovations. One possible explanation is that firms with significant global connection tend to have higher foreign ownership together with more overseas customers and suppliers too.

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]
Firm characteristics					
<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]
CEO characteristics					
<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]
Dummies					
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 the parenthesis. *, ** and *** represent statistical significance at the 10%, 5%, and 1% level. Constant term is included in all regression, but the results are not reported to save space. Panel B reports the marginal effect of Probit estimation.

tern emerges, suggesting that the positive association cannot be driven by the use of the linear probability model.

While experimenting different sets of controls in Columns 3-5, both panels show that there are rather small changes in the size of the estimates. We interpret it as an important evidence that while completely ruling out omitted variable bias is impossible, they are unlikely to be the major driving forces for the positive association.

Another cross-check of our result is to compare the economic significance of our estimate with that of Gorodnichenko, Svejnar and Terrell (2010). A one standard deviation increase in *Import penetration* corresponds to around 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.¹⁵

4.1 Instrumental variable estimation

Although Table 2 alleviates our concern for the omitted variable bias, endogeneity may come from potential reverse causality too: markets more flooded with foreign imports might be a luxury that only more innovative firms can afford to stay. In addition, as discussed, the noisy measure of firm-level import penetration results in measurement errors that biases down the estimates. We therefore use an IV. 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

¹⁵The 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 $\frac{0.12}{0.693}0.259 = 4.48\%$.

variables due to random-sampling of firms. But there can be a variety of industry-city level shocks that may affect industry-city-level innovation ($Innovation_{ic}$) and firm-level innovation ($Innovation_{fic}$). If these shocks are somehow picked up 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. Conditioning on industry and city dummies, our IV is therefore negatively correlated with imp_{fic} .¹⁶

The identifying assumption is that conditional on our controls, the IV does not affect innovation through channels other than the import penetration faced by the concern firm, i.e.,

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

This assumption is plausible because the industry-city-level shocks that affect innovation mentioned above is less likely to be picked up by import penetration faced by firms in other cities. In addition, while the measurement errors can be potentially serious at the firm-level, the average of firms' import penetration likely averages out firms' measurement errors. The

¹⁶Intuitively, 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 to one industry, the lower is its total imports. As the industry dummies control for the industrial policies, the inter-city difference within an industry reflects the differential impacts of the industrial policies across different cities. Thus, given the total amount of imports, if the policies have more impact on one city, it would mean having lesser impact in another city.

IV is therefore unlikely to systematically correlate with any firm-level measurement errors.

Table 3: Instrumental variable estimation

	1	2
Panel A: 2nd stage of 2SLS	Dependent Variable: Innovation	
Import penetration	0.802**	0.567**
	[0.356]	[0.271]
Firm characteristics	No	Yes
CEO characteristics	No	Yes
Dummies		
Competition dummies	No	Yes
Industry Dummies	Yes	Yes
City Dummies	Yes	Yes
Panel B: 1st stage of 2-step GMM IV	Dependent Variable: Import penetration	
Import penetration _{<i>i</i>,<i>c</i>}	-4.667***	-5.822***
	[0.645]	[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***	0.799***
	[0.343]	[0.314]

Note: White-robust standard errors clustered at the industry-city level are reported in the parenthesis. *, ** and *** represent statistical significance at the 10%, 5%, and 1% level. Constant term is included in all regression, but the results are not reported to save space. The first stages of the two-step GMM 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.

Panel B of Table 3 shows that indeed the IV is negatively and statistically significantly correlated with the endogenous variable. Other statistics help rule out the concern of weak instruments.¹⁷ Panel A shows that *Import penetration*, after being instrumented, still associate positively and significantly with *Innovation*. Panel C shows the corresponding marginal effect from IV Probit estimation. These results provide further confidence on the results shown in the OLS estimation, that imports do spur incremental innovation.

The IV estimates are larger than the corresponding OLS estimates. The increase in size confirms our concern that there can be potentially serious measurement errors on *Import penetration*, may bias the OLS estimates upward. However, it is also possible that our IV correlates positively with the error term or in other words, the failure of the orthogonal condition (5), which bias up the IV estimates. In the following we conduct two tests to support our claim that the strong association between import penetration and incremental

¹⁷The F-statistic is significantly above the critical value (10) of the “safe zone” for strong instrument (Staiger and Stock, 1997).

innovation is unlikely driven by the mis-use of an invalid IV.

4.1.1 A Falsification Test

For a check on the IV's orthogonal condition, Angrist and Pischke (2009) suggest a falsification test: if some variables are not supposed to be affected by the endogenous variable (e.g., pre-determined variables), a reduced-form regression of those variable on the IV should result in no significant association.

The SCE has a question asking the firm "According to your tax reporting requirements do you have to use a cash register or other electronic devices?" While tax reporting can itself be innovative too, a more innovative firm would not choose a system that is not required by the tax authority. As well, it is implausible to argue that the Chinese tax authority take import penetration (of firms in other cities) into account in deciding which tax reporting devices to use. Conditioning on the controls, regressing the choice of using cash register/electronic devices on our IV should result in no significant correlation. Indeed, our IV is highly insignificant, with a p -value of 0.87.¹⁸ There is no association between what tax-reporting devices to use and the IV.

4.1.2 Imperfect IV

Does the significantly positive association between import penetration and incremental innovation hinge on the IV's orthogonal condition? We use the technique advanced by Nevo and Rosen (2011) to address this concern. They innovate a technique of using an imperfect IV to draw inference; the technique admits relaxing the orthogonal condition of the instrument (5), resulting in bounds of the parameter of interest. Assume a much *weaker* assumption that the IV may be endogenous too, but it is correlated with the error term in the same direction as the endogenous variable (Assumption 3 in Nevo and Rosen, 2011) and the IV is less correlated with the error term than the endogenous variable (Assumption 4 in Nevo

¹⁸The results are not shown but are available upon request.

and Rosen, 2011).¹⁹

If more innovated firms tend more likely to stay in markets with more imports than less innovated firms do, 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 instrumental variable 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, it is implausible to argue that our IV (industry-city level) is even more endogenous than our endogenous variable (firm-level). 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.²⁰

Define λ as the ratio between the correlation between the IV and the error term (denoted $\rho_{z\varepsilon}$) 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 λ , we can construct a “perfect” IV, denoted $v(\lambda)$, that is correlated with the endogenous variable but not the error term using this formula, $v(\lambda) = \sigma_x z - \lambda \sigma_z x$, where x and z are the endogenous variable and the IV, respectively, and σ_x and σ_z are their respective standard deviations.²¹ The problem, however, is that we do not know the value of λ . Assumption 4, that the endogenous variable is at least as endogenous as the IV, implies $\lambda \in [0, 1]$. 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 β .

Table 4 shows using the original IV, *Import penetration*_{*i,-c*}, the estimate is 0.567, while

¹⁹Assumption 1 and 2 are linear specification and exogenous variables for other controls.

²⁰Precisely, 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} , σ_x and σ_z are the standard deviations of x and z , respectively. Since both σ_x , σ_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.

²¹The proof of this claim is not reported but is available upon request.

Table 4: Bounding parameters: relaxing exclusion restriction

	1	2
Panel A: 2nd stage of 2-step GMM IV	Dependent Variable: Innovation	
IV	Original IV	Constructed IV
Import penetration	0.567** [0.271]	0.208** [0.083]
Firm characteristics	Yes	Yes
CEO characteristics	Yes	Yes
Dummies		
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 the parenthesis. ** represents statistical significance at the 5% level. Constant term is included in all regression, but the results are not reported to save space. The first stages of the two-step GMM estimations include the same set of control variables as in the corresponding second stage.

using the constructed IV taking $\lambda = 1$, the estimate is 0.208 and is statistically significant. In other words, β 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 results that imports spur incremental innovation remain robust.

4.2 Robustness

Technologically capable firms clustering together? Are our results driven by technologically capable firms within an industry clustering at where imports are high?

Table 5 presents conflicting evidence on whether technologically capable firms do so. Taking the industry-city level average (subscript i, c), *Import penetration* does correlate positively with whether any patent is granted in the past 3 years (*Patent*), and whether any engineer or technician is hired in 2002 (*Engineers & technicians*), but correlates negatively with that of whether any R&D personnels is hired in the past 3 years (*R&D personnels*).²²

Table 6 directly controls for these three industry-city level measures. Since *Import penetration* remain statistically significantly positive, it is unlikely that even if technologically capable firms may cluster at where imports are high, such a tendency is sufficient to drive

²²We also control directly these three firm-specific variables directly in the estimation. The results are almost the same. The results are not reported but are available upon request.

our result.

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.

Alternative measures. Table 7 shows our results are robust to alternative measures. As in Lin, Lin, and Song (2010), Column 1 uses a dummy variable (denoted by $R\&D$) that takes a value of 1 if the R&D expenditure in 2002 is positive and 0 if otherwise. 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* $_{i,c}$; it is better at averaging out any errors if firms mis-reports their own import penetration ratios.

Sample bias and attrition. While 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 8 contains a sub-sample of firms whose main business contributes more than 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 (IPW) 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.

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>i,c</i>}	-0.135 [0.146]		
Engineering & technicians _{<i>i,c</i>}		0.264** [0.112]	
R&D personnels _{<i>i,c</i>}			0.155** [0.072]
Firm characteristics	Yes	Yes	Yes
CEO characteristics	Yes	Yes	Yes
Dummies			
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 the parenthesis. *, ** and *** represent statistical significance at the 10%, 5%, and 1% level. Constant term is included in all regression, but the results are not reported to save space. The first stages of the two-step GMM estimations include the same set of control variables as in the corresponding second stage.

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 _{<i>i,c</i>}			0.400*** [0.137]
Firm characteristics	Yes	Yes	Yes
CEO characteristics	Yes	Yes	Yes
Dummies			
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 the parenthesis. *, ** and *** represent statistical significance at the 10%, 5%, and 1% level. Constant term is included in all regression, but the results are not reported to save space. Subscript *i, c* imply *Import penetration* is calculate as its average of all firms within an industry-city unit.

Table 8: Robustness check: Sample bias and sample attrition

	1	2	3
Dependent variable: Innovation			
Import penetration	0.153*	0.193**	0.186**
	[0.091]	[0.084]	[0.090]
Probability (reporting import penetration)		0	
		[0.000]	
Probability (reporting import penetration) ²		-0.034*	
		[0.018]	
Firm characteristics	Yes	Yes	Yes
CEO characteristics	Yes	Yes	Yes
Dummies			
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 the parenthesis. *, ** and *** represent statistical significance at the 10%, 5%, and 1% level. Constant term is included in all regression, but the results are not reported to save space. 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.

5 Is competitive effect an underlying mechanism?

Firms may strive to better their products/services 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? According to Aghion, Bloom, Blundell, Griffith and Howitt (2005), this underlying mechanism implies that increased import competition reduces the pre-innovation rents more than it reduces post-innovation rents. The competitive effect should be weaker if foreign competition cannot easily vary pre-innovation rents. We exploit this idea with three strategies to indirectly test whether competition effect is an underlying mechanism.

5.1 Heterogeneous Response Estimation

In China, the trading relationships with governmental agencies are largely determined by political connection rather than economic considerations. These sticky relationships mean the pre-innovation rents of firms with more sales to government agencies should be less

influenced by increased import competition. With the share of output sold to government agencies (denoted by *State sales*), we use the strategy in Rajan and Zingales (1998).

Table 9: 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]
Firm characteristics	Yes
CEO characteristics	Yes
Dummies	
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 the parenthesis. *, ** and *** represent statistical significance at the 10%, 5%, and 1% level. Constant term is included in the regression.

Table 9 shows that the interaction term between *Import penetration* and *State sales* has a negative and statistically significant estimation 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 check whether firm-level technological capability is strongly correlated with the measure of governmental sales. Table 10 shows that there is indeed no strong evidence.²³

Table 10: 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.

²³We also regress *State sales* with each of the three technological capability measures, with and without controlling for all the other controls. The results are that there is no unified relationship. The results are not reported but are available upon request.

Table 11: When do imports matter? Size dimension

Dependent variable: Innovation	
Import penetration	0.289** [0.114]
Medium player	0.035 [0.033]
Big player	0.083*** [0.031]
Import penetration x Medium player	0.009 [0.199]
Import penetration x Big player	-0.313* [0.174]
Firm characteristics	Yes
CEO characteristics	Yes
Dummies	
Competition dummies	Yes
Industry Dummies	Yes
City Dummies	Yes
Number of Observations	1216
R-squared	0.1849

Note: White-robust standard errors clustered at the industry-city level are reported in the parenthesis. *, ** and *** represent statistical significance at the 10%, 5%, and 1% level. Constant term is included in the regression.

5.2 Dominant versus non-dominant firms

Table 11 includes dummies for big and medium players (versus small players), as well as their interaction terms with *Import penetration*.²⁴ The idea is that firms with a dominant market share are more likely to possess inherent attributes that is relatively hard to be substituted by imports, relative to firms without dominance. Therefore, their pre-innovation rents should respond less to increased import competition.

First, *Import penetration* is again positive and significant. Second, big players do seem to be more innovative, which may be a reason why they have gained dominance. Third, while the interaction term between *Import penetration* and *Medium player* is not statistically significant, that between *Big player* is negative and statistically significant. Imports are more likely to affect incremental innovations for small and medium players but not for big players.²⁵

²⁴The grouping depends on the firm's market share: small players (with market share less than 10%), medium players (with market share between 10% and 40%), and big players (with market share larger than 40%).

²⁵The results are qualitatively similar when we use other cutoff points, such as 5% for medium player and 50% for big player.

5.3 Manufacturing versus services firms

Manufacturing outputs are generally more trade-able than services. This creates a natural shield for firms in the services sector from foreign competition. Columns 1-2 of Table 12 contrast the manufacturing and service firms. Though both estimated coefficients are positive, only that of manufacturing firms is significant. The much smaller sample size may be a reason why that of service firms is marginally insignificant.

Columns 3-4 follow 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, chemical products and medicine, and biotech products and Chinese medicine).²⁶ The discouraging effect of Aghion, Bloom, Blundell, Griffith and Howitt (2005) dominates if the laggard firms are further away from the technological frontier. We find consistent pattern that *Import penetration* is positive and significant only 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 their imports are robust determinants of their incremental innovations. This echoes well with the finding of 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.

²⁶As the industrial classifications in China are different from those in the US, 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.

Table 12: When do imports matter? Industry dimension

Sample	1	2	3	4
	Manufacturing firms	Services firms	Mature industries	High-tech industries
	Dependent variable: Innovation			
Import penetration	0.154* [0.091]	0.327 [0.199]	0.267*** [0.100]	-0.044 [0.152]
Firm characteristics		Yes	Yes	Yes
CEO characteristics		Yes	Yes	Yes
Dummies				
Competition dummies	Yes	Yes	Yes	Yes
Industry Dummies	Yes	Yes	Yes	Yes
City Dummies	Yes	Yes	Yes	Yes
Number of Observations	913	317	525	388
R-squared	0.188	0.2017	0.2088	0.19

Note: White-robust standard errors clustered at the industry-city level are reported in the parenthesis. *, ** and *** represent statistical significance at the 10%, 5%, and 1% level. Constant term is included in all regression.

Our paper focuses on ruling out other alternative explanations. Using different strategies together with a battery of robustness checks, our paper documents that the result is unlikely to be due to other global connection, 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 better their products, an underlying mechanism for understanding the responses of firms in the North.

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] **Angrist, Joshua D., and Jorn-Steffen Pischke.** 2009. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press, Princeton, NJ.
- [3] **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*.

- [4] **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*.
- [5] **Bloom, Nicholas, Paul Romer, and John Van Reenen.** 2010. "A Trapped Factor Model of Innovation." *Working paper*.
- [6] **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.
- [7] **Dong, Xiao-yuan, and Lixin Colin Xu.** 2009. "Labor Restructuring in China: Toward a Functioning Labor Market." *Journal of Comparative Economics*, 37(2): 287-305.
- [8] **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.
- [9] **Gorodnichenko, Yuriy, Jan Svejnar, and Katherine Terrell.** 2010. "Globalization and Innovation in Emerging Markets." *American Economic Journal: Macroeconomics*, 2(2): 194-226.
- [10] **Henderson, Vernon, Ari Kuncoro, and Mathew Turner.** 1995. "Industrial Development in Cities." *Journal of Political Economy*, 103(5): 1067-90.
- [11] **Hausman, Jerry, Gregory Leonard and J. Douglas Zona.** 1994. "Competitive Analysis with Differentiated Products." *Annales D'Economie Et De Statistique*, 34: 159-180.
- [12] **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).

- [13] **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.
- [14] **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.
- [15] **Lin, Ping, and Kamal Saggi.** 2002. “Product Differentiation, Process R&D, and the Nature of Market Competition.” *European Economic Review*, 46(1): 201-211.
- [16] **Nevo, Aviv, and Adam Rosen.** 2011. “Identification with Imperfect Instruments.” *Review of Economics and Statistics*, forthcoming.
- [17] **OECD.** 2008. *Science, Technology and Industry, 2008*. Paris: Organisation for Economic Co-operation and Development Press.
- [18] **OECD.** 2006. “South-South Trade: Vital for Development.” *Policy Brief*, August 2006. Paris: Organisation for Economic Co-operation and Development Press.
- [19] **Rajan, Raghuram G., and Luigi Zingales.** 1998 “Financial Dependence and Growth.” *American Economic Review*, 88(3): 559-86.
- [20] **Rodrik, Dani.** 2006. “What’s So Special About China’s Exports?” *China and World Economy*, 14(5): 1-19.
- [21] **Schott, Peter K.** 2008. “The Relative Sophistication of Chinese Exports.” *Economic Policy* 23(53):5-49.
- [22] **Staiger, Douglas, and James H. Stock.** 1997. “Instrumental Variables Regression with Weak Instruments.” *Econometrica*, 65(3): 557-586.
- [23] **Teshima, Kensuke.** 2009. “Import Competition and Innovation at the Plant Level: Evidence from Mexico.” *Working Papers*.

- [24] **UNCTAD.** 2005. *Trade and Development Report 2005*. United Nations publication, New York and Geneva.
- [25] **Wang, Zhi, and Shang-Jin Wei.** 2010. “What Accounts for the Rising Sophistication of Chinas Exports?” In *China’s Growing Role in World Trade*, edited by Robert C. Feenstra and Shang-Jin Wei, 63-108. Chicago: University of Chicago Press.
- [26] **Williamson, Peter J., and Ming Zeng.** 2007. *Dragons at Your Door: How Chinese Cost Innovation Is Disrupting Global Competition*. Cambridge, Massachusetts: Harvard Business School Press.
- [27] **Wooldridge, Jeffrey M.** 2002. “Inverse Probability Weighted M-estimators for Sample Selection, Attrition, and Stratification.” *Portuguese Economic Journal*, 1(2): 117-139.
- [28] **Wooldridge, Jeffrey M.** 2007. “Inverse Probability Weighted M-Estimation for General Missing Data Problems.” *Journal of Econometrics*, 141(2): 1281-1301.
- [29] **World Trade Organization.** 2005. “Developing countries’ goods trade share surges to 50-year peak,” April 14, Press Release/401.

A Data appendix

Cull and Xu (2005), Dong and Xu (2009), and Lin, Lin, and Song (2010), among others, use the dataset too. Meanwhile, there is a new survey of Chinese enterprise recently conducted by the World Bank covering firms in 100+ cities in China (the third wave in 2005). However, that dataset does not include many variables present in this early dataset.

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 the survey year of 2002, while the second part pertains to the period of 2000-2002. They 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).