Offshoring, R&D and Productivity Growth: Evidence from U.S. Census Microdata*

Ishan D. Ghosh†
Drexel University
idg23@drexel.edu

Job Market Paper

August 16, 2019

Abstract

This paper studies the impact of foreign input sourcing on R&D expenditures and quantifies their complementary effect on firm performance. Using confidential firm-level microdata from the U.S. Census Bureau, I first use a shift-share designed instrument to exploit exogenous variation in offshoring and show that foreign intermediate sourcing is associated with increased domestic R&D expenditures. Second, I build and estimate a structural dynamic model of R&D investment, in which a decline in the relative cost of imported intermediates leads to an increase in R&D investment and thus endogenously, an increase in firm productivity. I then use the estimated model to quantify the effects of a (not so) counterfactual increase in price adjusted quality advantage of foreign intermediates and find that firm productivity growth is positively affected, partly due to the increase in static gains from offshoring and partly due to the resultant higher endogenous participation in R&D. Structurally estimated decline in the relative price of imported intermediates has a significant positive impact on firm performance equivalent in magnitude to that of a 25% R&D cost subsidy.

JEL classification numbers: L25,F14,F61
Keywords: Offshoring, R&D, Productivity, International Trade, Firm Dynamics

---

*I am extremely grateful to advisor André Kurmann for his invaluable advice, guidance and continued support, and to my committee Yoto Yotov, Phillip Luck, Mian Dai and Konstantinos Serfes. I thank Fariha Kamal, Hongsong Zhang, Ben Hyman, Justin Pierce, Teresa Fort, Matthias Kehrig, Nik Zolas, Chen Yeh, Ben Pugsley, Kyle Handley, Erika McEntarfer, Alejandro Molnar, Kyle Herkenhoff, Kamran Bilir for valuable discussions and comments. I thank conference participants at Comparative Analysis of Enterprise Data (CAED), GCER 2019 Biennial Conference, Washington, DC, Kauffman Foundation Entrepreneurship Bootcamp, 2019. Any opinions and conclusions expressed herein are those of the author(s) and do not necessarily represent the views of the U.S. Census Bureau. All results have been reviewed to ensure that no confidential information is disclosed

†School of Economics, LeBow College of Business, Drexel University. www.ishanghosh.com

1 Introduction

Do trade barriers affect domestic innovation and performance of firms? With the rise of China as a major trading partner and the integration of the global economy in general, a burgeoning literature focuses on the impact of trade on U.S. labor market dynamics, as low-skilled, routine jobs move out of developed countries. On the other hand, there is growing interest in how firms engaging in trade are able to increase their scale of operations and profitability by fragmenting production across multiple offshore locations and retaining stages of core competencies within the headquarter country or expanding in service sector jobs (Bernard et al. (2018), Bloom et al. (2019)). However, little is known on the exact channels through which firms are able to take advantage of offshore production cost differentials to augment their performance. In light of the current political discourse on protectionist trade policies, research on these channels (or the absence thereof) is fundamental to quantify the consequences of trade policies on firm performance.

In this paper, I quantify the effect of access to cheaper intermediate inputs through offshoring on R&D expenditures and their joint impact on firm productivity growth. Complementarity in offshoring and R&D arises on account of reduced production costs and resultant increased profitability due to offshoring, which augments firms' propensity of engaging in R&D, thereby endogenously affecting firm performance. Hence, in addition to affecting labor market outcomes, increased trade barriers could have a long lasting effect on relocating innovation away from the domestic economy, which can result in lower average productivity growth. In order to empirically analyze the offshoring-R&D complementarity, I combine confidential microdata from the U.S. Census Bureau's economic censuses, trade transactions, and innovation surveys into a new firm-level database spanning nearly two decades from 1997 to 2015. This database contains detailed shipments, production inputs and imports of the majority of large firms with manufacturing establishments in the U.S., along with their levels of R&D expenditures and earnings over time. To the best of my knowledge, this paper is the first to understand the productivity and innovation effects of offshoring with U.S. administrative microdata.

Using this database, I, first, start by performing a reduced-form analysis, where I relate time varying firm level measures of offshoring to R&D outcomes. Second, I build and estimate a forward-looking structural dynamic model of R&D investment in which a decline in the relative cost of imported intermediates leads to an increase in R&D investment and endogenously, an increase in firm productivity. I then use the estimated model to quantify the effects of a (not so) counterfactual increase in tariffs on foreign intermediates and find that firm performance decreases partly due to decline in static gains from offshoring and partly due to the resultant lower endogenous participation of R&D.

The reduced form analysis builds on the strategy of Hummels et al. (2014) which identifies exogenous variation in firm offshoring due to increased productivity of trading partners, apportioned to firms based on their pre-sample country-product[1] specific import links. The exclusion restriction requires that trading partners' improved productivity proxied by their product specific export growth to countries other than the U.S., affects R&D outcomes only through offshoring. However, firms in the U.S. could adjust their R&D engagement in response of import competition as argued by Bloom, Draca and Van Reenen (2015) and (Autor et al., 2019). I control for such effects using detailed industry by year fixed effects to absorb any time varying industry shocks like import competition.

Several important findings stand out from the reduced form analysis. *First*, controlling for various technology augmenting firm level covariates like size, sales, capital stock and labor productivity, I find that offshoring has a positive and significant effect on R&D expenditures. On average, a 10% increase in the level of offshoring leads to more than 0.7% increase in R&D expenditures.[2] In terms of economic

[1]Product here refer to HS6 product classifications.

[2]The baseline measure of offshoring is the sum of all imports industries that are similar upto the first **four** digits of the

significance, for firms experiencing more than 50% annual growth in offshoring (accounting for 22% firm-year observations), the median R&D expenditure growth is approximately 23%. *Second*, while offshoring has no effect on average manufacturing earnings, the effect on R&D earnings is strongly positive and statistically significant at the 1% level.[3] *Third*, conditional on scale effects, the effect of offshoring on R&D expenditures and earnings monotonically increases with the degree of similarity between imported products and domestic shipments, thereby implying greater value added to R&D in response to increased substitution of domestic labor with foreign resources.[4]

To understand the joint effect of offshoring and R&D on firm performance, I introduce foreign intermediates in a structural model of endogenous productivity following Halpern, Koren and Szeidl (2015) and estimate the model in two sub-samples, 1997-2006 and 2007-2015. The model estimates significant *static* gains from importing, stemming primarily from access to a wider *variety* of intermediates. Unlike prior evidence from developing countries like Hungary and Colombia, there is no evidence of the *quality* channel as U.S. firms on average import price-adjusted lower quality intermediates, especially during 1997-2006. However, consistent with Fan, Li and Yeaple (2015), I find the average quality effect to increase substantially in the second half of the sample, when imported intermediate inputs have no obvious quality disadvantage after adjusting for price difference. Offshoring, on average increases current revenue by 31%[5], while undertaking R&D increases average within-firm productivity in the next period by 6%. On one hand, this positive productivity affects future periods through the Markov evolution process, while on the other hand, increased productivity enables firms to self-select into R&D which further augments future productivity.

Addressing the endogeneity in R&D choice, I estimate a forward-looking dynamic structural model of R&D investment choice, where it depends on the expected future profits and current fixed or sunk costs of performing R&D. The dynamic model reinforces that productivity is endogenous, being positively affected by R&D and that entry costs are a significant factor in firms' decision to undertake R&D.[6] The dynamic structural model allows me to estimate R&D sunk and fixed costs, which on average range from $603,197 - $1,077,323 and $43,477 - $51,534 respectively. I then use estimated model parameters to conduct a policy experiment and evaluate how declining trade barriers increases firm value via offshoring (*static effect*) which endogenously affects R&D (*dynamic effect*). To that end, I examine the impact on firm performance in the first sub-period, by setting the price adjusted quality parameter to its estimated value in the second sub-period, to model a decline in relative cost of intermediates. The policy experiment highlights a significant 7.75% increase in average firm value, of which 28.21% is due to increase in static gains from offshoring, while the remaining 71.78% is attributed to the dynamic effect, owing to increased participation in R&D activities.

To provide economic context to the policy experiment, I compare the gain in average firm performance to a second counterfactual where I decrease R&D fixed and sunk costs by 25%. Reduction in R&D entry barriers increases average firm value (only via the dynamic effect). Comparing between the two policy experiments, I show that the effect of a structurally estimated decline in the relative price of imported intermediates is equivalent in magnitude to that of a 25% R&D cost subsidy. R&D cost subsidies are considered to be one of the primary drivers of fostering innovation. These counterfactual experiments show that spillovers from imposing trade barriers could have a similar effect as that of R&D taxes, thereby

shipment industries.

[3]To construct average R&D earnings, I divide total R&D payroll by the number of R&D workers reported by the firm.

[4]I use four versions in total - imported inputs in the same 3 digit/4 digit/5 digit and 6 digit NAICS industry as that of the firm's shipments. The most restricted version which required the firm to import the exact six digit NAICS product(s) as their shipment(s), is closest to the measure used by Bernard et al. (2018).

[5]This estimate is higher than 20% as inBoler, Moxnes and Ulltveit-Moe (2015) and 22% as in Halpern, Koren and Szeidl (2015). Both papers use a different measure of offshoring based on the *number* of imported products. In Appendix section XX, I use a similar measure and obtain identical estimates.

[6]Unlike Aw, Roberts and Xu (2011), I do not find firms' export choice to have a significant effect on future productivity.

dampening innovation in the domestic economy.

Related Literature: This paper brings together a few different, burgeoning strands of literature on the intersection of international trade, firm dynamics and industrial organization. *First*, this paper relates to the literature on the effect of intermediate sourcing on firm productivity in different countries like Indonesia (Amiti and Konings (2007)), Chile (Kasahara and Rodrigue (2008)), Hungary (Halpern, Koren and Szeidl (2015)) and India (Topalova and Khandelwal (2011)) to name a significant few. In estimating the static gains from importing, my paper is closest to Halpern, Koren and Szeidl (2015), who use a structural approach to estimate two different mechanisms of import gain for a panel of Hungarian firms: a quality and a variety channel. Halpern, Koren and Szeidl (2015) find that importing inputs increases firm productivity by 22%, one-half of which is attributed to an increase in the variety of intermediates used in production.[7] I estimate a similar structural model to find a similar variety effect but find no evidence of the quality channel. This seems reasonable as U.S. firms, unlike Hungarian firms, do not necessarily import higher quality intermediates from low wage countries.

Second, this paper is related to a relatively new literature on trade and innovation. With a focus on import competition, there is conflicting evidence on the effect of Chinese imports on innovation. While Bloom, Draca and Van Reenen (2015) using data from OECD countries find evidence of increase in R&D, patenting and TFP, Autor et al. (2019) find that, among U.S. firms, patenting declined in response to the escalating threat of Chinese import competition. While Bloom, Draca and Van Reenen (2015) and Autor et al. (2019) look at firm responses to industry-level shocks increasing competition, this paper focuses on the effects of foreign inputs sourced by firms themselves. To that end, the reduced-form approach in my paper is very similar to Bernard et al. (2018), who focus on firm imported inputs and their effect on within-firm skill composition. While Bernard et al. (2018) attempt to suggest a link between offshoring and change in domestic skill composition, my overall focus is to quantity the effect of trade on firm performance via innovation.

Third, this paper is very closely related to a small but constantly evolving, literature on the interdependent role of trade and R&D on firm productivity. Building on Aw, Roberts and Xu (2011), Boler, Moxnes and Ulltveit-Moe (2015) was the first paper to explore the joint complementarity of R&D and imported intermediates on firm productivity. Boler, Moxnes and Ulltveit-Moe (2015) study the impact of a R&D cost shock on R&D investment, imported inputs and their joint impact on firm productivity in Norway. The key to their channel was the exogenous tax policy on R&D for a subset of firms which incentivized firms to not only undertake more R&D expenditure, but also expand internationally by sourcing cheaper inputs. I focus on an alternate channel which depends on exogenous changes in trade costs to spur R&D and in turn boost firm performance. To the best of my knowledge, this paper is the first to quantify the trade driven R&D complementary effect on firm productivity. Such complementarities, in the wake of current policy debates on trade barriers, have an enormous role to play in highlighting gains from trade on the domestic economy via innovation and productivity growth of firms.

Overview of this paper. The remainder of the paper is as follows. Section 2 describes the data while Section 3 describes measurement issues and endogeneity, along with stylized facts on offshoring to motivate the empirical analysis. Section 4 outlines the reduced-form empirical strategy and results. Section 5 develops and estimates the model. In Section 6 I present a counterfactual exercise, to quantify the effect and relative importance of the proposed theoretical mechanism, and Section 7 concludes.

[7] The other component of import gain is attributed to increased access of higher quality intermediates.

2 Data Sources

The results of this paper are based on linking various U.S. Census Bureau's restricted use micro datasets. I use trade data from the Longitudinal Firm Trade Transactions Database (LFTTD), manufacturing establishments' input and output data from the Annual Survey of Manufactures (ASM) and the Census of Manufactures (CM), data on R&D expenditure and employment from the Survey of Industrial Research and R&D (SIRD) and Business R&D and Innovation Survey (BRDIS). Ultimately all micro datasets are linked together using the Longitudinal Business Database (LBD) firm (*firmid*) and establishment (*lbdnum*) identifiers. In what follows, I briefly describe how I use the different datasets for the analysis, while detailed descriptions of the data and linking processes are relegated to the Appendix (sections A.XX to A.XX).

This analysis requires firms to have some manufacturing presence in the U.S., to potentially substitute domestic production with increased input sourcing from abroad. Aggregating establishment level data in the LBD to the firm, I restrict the sample to roughly 30,000 firms with positive manufacturing employment in a given year. For the given sample of firms, I obtain information on average production earnings, number of production workers, and shipments by detailed product codes, from the ASM and CM along with their product trailer files. In order to study firm level import of intermediates, I exploit the LFTTD, which contains import and export values (nominal US$) for the universe of trade transactions over time by firm, product (HS 10 digit) and source country. The LFTTD is then merged to the sample of manufacturing firms[8] using the longitudinally consistent and distinct *firmid* variable.

Finally, I obtain R&D data from two surveys - SIRD (1997-2007) and BRDIS (2008-2015) that collect annual data on firms' R&D expenditures and personnel, both within the U.S. and overseas. The SIRD and BRDIS are the only sources of R&D data for both private and publicly listed firms[9] and proves to be an invaluable source to obtain a representative sample of R&D performers in the economy.

3 Measuring Offshoring

I begin this section with a discussion on how I define offshoring and how it differs from other measures previously used in the literature. I then discuss new stylized facts about offshoring and innovation and how they relate to aggregate trends in the data. It is important to note that offshoring might systematically differ from total firm level imports of products. The key to the analysis of offshoring is whether the observed firm-level imports of products are final goods or intermediates in production and whether the firm sources these inputs with an intent to substitute domestic production with foreign resources. The LFTTD trade data does not distinguish final from intermediate imports and nor does it provide information on how the firm uses these imported products. Using input-output coefficients as weights, Feenstra and Hanson (1999) defined "narrow offshoring" as purchases of inputs belonging to the same industry as that of producing firms.[10]

Hummels et al. (2014) applies this idea to individual firms by defining the "narrow measure" of offshoring to be the sum of imports in the same HS4 category of goods sold by the firm (either domestically or via exports). Although I follow Hummels et al. (2014) very closely, I use a more flexible definition, taking into account fundamental differences in data and research design. The combination of shipped and imported products is essential to the construction of the firm level measure of offshoring. Instead of restricting imports in the same HS4 category, I aggregate both trade and shipment product data to six digit NAICS

[8] In-sample firms with positive manufacturing employment are referred to as manufacturing firms.

[9] Compustat provides R&D data for only publicly listed firms.

[10] Imports of computer microchips by the electronics industry would be classified as narrow offshoring, but those same imports by the automobile industry would not (Hummels et al. (2014).

classification,[11] and then sum imports which are in the same three, four, five and six digit NAICS codes as that of that manufactured products. This allows substantial flexibility in measuring offshoring and also tests if more restricted imports, implying a higher degree of similarity, has a significantly differential impact on innovation outcomes.

Note, that this measure is a strict lower bound of firms' true offshoring intensity, since it only takes into account foreign intermediates sourced by the firm itself. It ignores the large wholesale market, whereby wholesale firms import upstream intermediates and sell them to domestic manufacturing firms. The domestic manufacturing firms benefit from the imported varieties and can consequently substitute foreign intermediates for domestically produced counterparts. Ganapati (2017) finds that between 1997 and 2007, the share of transactions intermediated by wholesalers increased 34%, with internationally sourced varieties accounting for half of this gain. Unfortunately, my data does not contain firm to firm shipments and hence I characterize offshoring only when the firm itself imports intermediates similar to its shipped varieties.[12]

3.1 Stylized Facts

Next, I present stylized facts in the data, which highlights three main ideas that help to motivate the empirical analysis: (i) "Narrow offshoring" is an important dimension in firm imports, (ii) Offshorers produce the same products that they import at a very narrowly defined industry classification, (iii) Offshorers import a wider variety of goods compared to traditional importers, (iv) Offshorers are larger, more productive and have a higher propensity to conduct R&D.

Fact 1 *"Narrow Offshoring" on average accounts for a significant share in firm imports.*
Table 1 shows that the different restricted versions of "narrow offshoring", on average accounts for a significant share in firm imports. My baseline measure[13] which includes imports in the first four (NAICS 4) digits of shipped industries account for approximately 64% of total imports[14].

Table 1: Share of (narrow) Offshoring in Imports

Level of Narrow Offshoring	Import Share
3 digit	0.718
4 digit	0.638
6 digit	0.538

Fact 2 *Offshorers produce the same good as they import.*
Similar to Bernard et al. (2018), I also observe that offshorers, instead of "hollowing out" their domestic production, continue to produce the same goods that they import. Final row of Table 1 shows that on average, the most restricted measure of definition which includes imports at the exact same six digit NAICS industry that the firm produces, accounts for more than 53% of imports. Unfortunately due to the lack

[11] The product structure in the LFTTD is at the HS 10 digit. I use the product-level concordance by Pierce and Schott, to map the HS 10 digit import flows into NAICS-6 digit industry codes. The CM and ASM follows a NAICS-10 digit product structure. In the absence of a clean mapping from HS-10 to NAICS-10, I aggregate the shipment data to six digit NAICS industry codes.

[12] Such data can be obtained with a certain degree of limitation on coverage from The Commodity Flow Survey (CFS), which is conducted every five years and collects data on a random selection of shipments for a set of establishments. This data is collected for both wholesale and manufacturing establishments and is used to construct crosswalks between manufacturing and wholesale sectoral designations.

[13] Four digit restricted offshoring is the preferred baseline throughout the paper. Offshorers (unless specified otherwise) refers to firms that have positive NAICS 4 digit restricted imports.

[14] Boehm, Flaaen and Pandalai-Nayar (2019) finds the share of manufacturing imports classified as intermediates in 2007 to be 64 percent.

of reliable quantity data, I am unable to test if imported varieties have lower unit prices (proxying for quality) compared to domestically produced varieties of the same product, similar to Bernard et al. (2018). Instead I rely on a structural model to estimate an industry specific effect, showing that price differences in imported and domestic intermediates reflect their quality difference on average.

Fact 3 *Offshorers import a wider variety of goods relative to non-offshoring importers.*
Table 2 shows that although firms on average concentrate their import spending on top two or five products[15], importing firms in my sample exhibit a somewhat different behavior. In my sample, importers (not necessarily offshorers) have a less concentrated import spending, while offshorers particularly import a wider variety of products as on average they spend 68% and 86% on top 2 and 5 products respectively. Other than the quality channel, importing firms in developed countries like the U.S. and Denmark, gain from a higher variety of intermediates even within narrowly defined industries. This is also consistent with a larger share of restricted imports in the same six digit narrow NAICS industry as the shipments of the firms, a firms do not completely replace domestic production and use offshoring to gain from the *variety effect*.

Table 2: Share of Top Products in Imports (By value)

Sample	Share of Top 5	Share of Top 2
LFTTD	0.974	0.902
Estimation Sample	0.89	0.734
Offshorers in Sample	0.861	0.682

Notes: Share of Top 5 and 2 shows the fraction of total imports accounted for by the top 2/5 products of a given firm-year.

Fact 4 *Offshorers are larger, perform more R&D and have higher sales*[16]
Similar to Boler, Moxnes and Ulltveit-Moe (2015), I run a set of simple regressions highlighting the difference between offshoring and non-offshoring firms.[17] Log firm characteristics like employment, R&D expenditures and shipments are used as outcome variables and a dummy variable indicating whether a firm engaged in offshoring or not.[18] Table 3 identifies the positive correlation between within firm offshoring and firm characteristics implying that firms expand employment, R&D and shipments, when they begin to offshore.

I now characterize the firms in my sample in terms of their trading and innovative activities and in the context of aggregate trends in the data over the sample period 1997 to 2015. Table 4 reports the importance of trade and innovation at the firm level. Owing to the general research question, the sample of firms involved in trade and R&D, inherently focuses on large firms as observed in the first panel of Table 4.

With an average firm size of around 400 employees, the low standard deviations indicate that these firms do not vary much in terms of their output and capital stock. However the standard deviations of the trade variables are considerably high, which allows me to exploit significant variation in offshoring across

[15] Hummels et al. (2014) shows that top 2 and 5 products account for 67.9% and 92.1% of total imports respectively in Denmark. U.S. firms also exhibit strikingly close shares. From row 1 of Table 2, I show that across all importing firms in the LFTTD, the share of top 2 and 5 products are 90.2% and 97.4 % respectively.

[16] I classify manufacturing shipments as sales. This might not be equal to total revenue earned by the firm.

[17] More formally I run: $\ln \mathrm{Y}_{it} = \alpha + \mathbb{1}\left[\beta_O * \mathrm{Off}_{it}\right] + \theta_i + \delta_t + \varepsilon_{it}$

[18] Once again, I use the baseline measure of offshoring i.e., restricted at the four digit NAICS as shipments. Hence offshoring = 1 if a firm has positive 4 digit restricted offshoring, 0 otherwise.

Table 3: Offshorer Premia

	Log Employment	Log R&D Exenditure	Log Shipment
Offshoring Dummy	0.143***	0.231***	0.192***
	(0.010)	(0.071)	(0.013)
Year fixed effects	✓	✓	✓
Firm fixed effects	✓	✓	✓
R^2	0.96	0.80	0.94
Observations	67000	67000	67000

Notes: Data used in this sample has firm-year observations from 1997-2915. Offshoring Dummy is a categorical variable which is equal to 1 for firms with positive offshoring (at 4 digits) and 0 otherwise.

Table 4: Descriptive Statistics

	Observations	**Mean**	**SD**
Firm Characteristics			
Log Employment	81500	5.977	1.752
Log Shipment	81500	18.24	1.922
Log Capital Stock	81500	16.9	2.075
Log Labor Productivity	81500	12.23	0.932
R&D			
R&D Dummy	81500	0.666	0.472
Log of R&D Expenses	81500	9.768	7.198
Earnings			
Log R&D Earnings	53000	8.298	4.882
Log Production Earnings	81000	3.601	0.369
Employment			
Log R&D Employment	81500	2.122	2.115
Log Production Employment	81500	5.193	1.679
Trade			
Log Offshoring (4 digits)	81500	8.918	7.774
Log Related Offshoring (4 digits)	81500	5.723	7.583
Log Offshoring (4 digits) > 0	48500	14.94	3.347
Number of Imported Products (HS4)	81500	22.59	39.25
Log Exports	81500	14.6	6.211

Notes: Data used in all panels has firm-year observations spanning the sample period 1997-2015 except for R&D earnings for which I use data from 1997-2013 due to concerns of data reliability. For each variable, I report the mean, standard deviation and the number of observations.

firm-years in the sample. The SIRD/BRDIS surveys are biased heavily in favor of R&D performers which is evident in the extensive margin of R&D performers. On average, 67% of firms engage in positive R&D across all years in the sample.

3.2 Instruments

In the reduced-form approach, I relate time varying innovation outcomes to time varying firm level measure of offshoring. Two key identification challenges exist. First, identification might suffer from firms engaging in R&D also engaging in offshoring, resulting in simultaneity bias. Second, unobserved productivity and/or demand shocks potentially correlated with firms' innovation outcomes, might introduce bias in the measurement of the causal impact of offshoring on R&D. As argued by Bernard et al. (2018), the direction of bias is unclear empirically. In the presence of unobserved negative productivity shocks like reduced demand or credit access, the OLS coefficients are likely to be biased downward, while positive productivity shocks will upward bias the OLS coefficient of offshoring.

To address these competing forces, I construct instruments that are correlated with the firms' level of offshoring but are uncorrelated with firms' innovation outcome. To that end, I use a standard shift-share design where I apportion changes in product-country specific world export supply (to destinations other than U.S.) to U.S. firms, based on their pre-sample product-country import shares. This shift-share design in the spirit of Hummels et al. (2014) and Bernard et al. (2018) enables me to identify changes in firm offshoring due to factors exogenous to the firm and U.S. in general. The world export supply identifies changes in product-specific comparative advantage of the exporting country due to factors relating to price, quality or economy-wide structural shocks. For example, Fan, Li and Yeaple (2015) finds that China's access to the WTO and subsequent reduction in tariffs induced Chinese exporters to upgrade the quality of their exports. More specifically, using country-industry level trade flows from COMTRADE, I construct the World Export Supply (WES) instrument for firm i, year t as:

$$I_{it} = \sum_{c,k} s_{ick} WES_{ckt} \tag{1}$$

where, $s_{ick} = \frac{IMP_{ick}}{\sum_{c,k} IMP_{ck}}$ represents the share of country c, product k imports as a share of firm i's total imports in the pre-sample year (*share*) and WES_{ckt} is export supply for country c in product k at time t (shift). The pre-sample weights are stable over time as Table 5 shows that roughly 62% of contemporaneous country(c)-product(k) specific import flows also appeared in the pre-sample, which is strikingly close to Hummels et al. (2014).[19]

The highly disaggregate trade data also contains information on whether imports were between related parties or sourced at arms-length. I use this information to decompose total offshoring into related-party and arms-length offshoring, and instrument for them accordingly: *type* $\in$ *(arms-length, related-party)*:

$$I_{it}^{type} = \sum_{c,k \in (type)} s_{ick}^{type} WES_{ckt} \tag{2}$$

where, $s_{ick}^{type} = \frac{IMP_{ick}^{type}}{\sum_{c,k} IMP_{ck}}$ represents the share of country c, product k related(arms-length) imports as a share of firm i's total imports in the pre-sample year (*share*). Table 5 shows that related party shares, (around 60%) slightly less in magnitude compared to total imports, are also relatively stable over time. [20]

Table 5: Pre-sample Flows

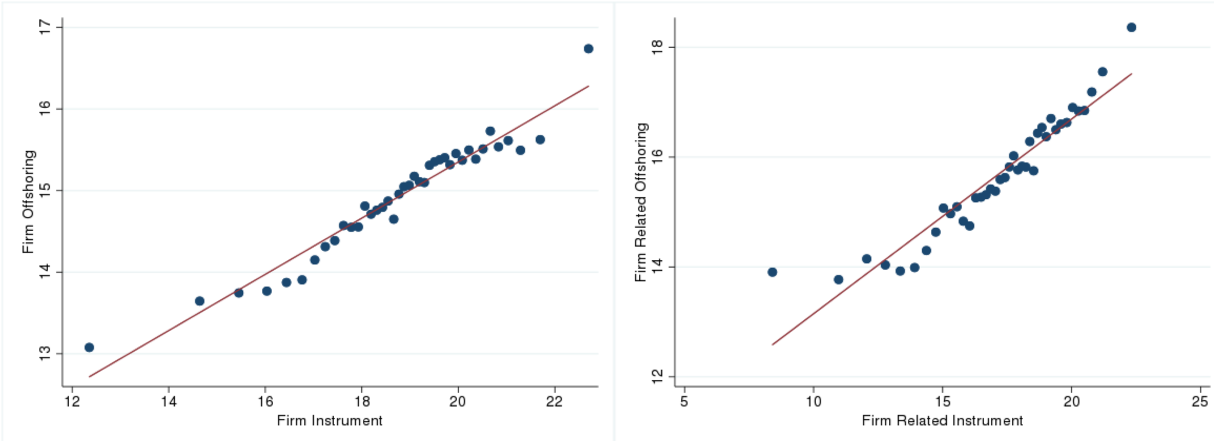
Imports	WES Weights	Share
Total	fpc	61.8
Related	fpc	59.7
Total	fp	79
Related	fp	76.1

Notes: Share refers to the fraction of total imports accounted for the the firm-product-country(fpc) and the firm-product(fp) flows for both total and related imports that appear in the pre-sample year (1996).

[19] Hummels et al. (2014) shows that 64.4 percent of c-k import flows purchased by firms in-sample also appeared in the pre-sample (conversely, roughly one-third of in-sample import purchases were not represented in the pre-sample).

[20] Instead of product-country specific weights, Bernard et al. (2018) uses product-specific weights which enables them to account for a higher share of in sample product-country imports. As a robustness check, I use China-specific offshoring and follow Bernard et al. (2018) to use product-specific weights to instrument for it.

Figure 1 sketches the estimation strategy. Panels 1a and 1b reveal substantial predictive power of the shift-share designed IV for total and related party offshoring respectively. The baseline IV uses country-product-firm weights to apportion changes in World Export Supply (WES) to alleviate concerns regarding insufficient product-share variation across firms within industries.



(a) Total Offshoring **(b) Related Party Offshoring**

Figure 1: First Stage Plots

3.3 Threats to Identification

In this section, I discuss potential threats concerning the validity of the shift-share design as instruments for offshoring as well as analyzing the direction of the bias in OLS. First, as argued by Hummels et al. (2014), rise in world export supply for a particular country-product pair can be induced not only by supply shocks, but also by overall world demand for that product k. I deal with this issue by including time varying fixed effects and by separately controlling for U.S. firms' exports, given the intuition that if demand for a particular product is worldwide, U.S. firms are likely to respond to such demand by exporting more of the given product.[21] Second, as argued by Bernard et al. (2018), the exclusion restriction for the instrument requires that changes in world export supply affects firm innovation outcomes only through offshoring. Using European firm-level data, Bloom, Draca and Van Reenen (2015) show that import competition can induce firms to upgrade technology in an attempt to increase productivity. I include detailed six digit NAICS industry by year fixed effects, which controls for potential time-varying industry demand and productivity shocks at the most disaggregate level in the data.

Even with the inclusion of detailed industry by time fixed effects and time invariant firm fixed effects, firms over time could respond differently to both domestic and foreign competition shocks, especially with respect to high cost investments like R&D. This residual effect of competition and productivity shocks over time is important to understand the direction of bias in OLS coefficients, particularly when these shocks are highly correlated with firms' decision and intensity to offshore production. In order to understand the mechanism, let us consider a simple econometric model where the dependent variable Y (for instance R&D) is affected by observable offshoring (X_1) and unobservable firm-specific import competition (X_2). Formally I can express this as:

$$Y_{it} = \beta_1 X_1 + \beta_2 X_2 + \varepsilon$$

[21] Results remain unchanged. Available upon request.

Since X_2 is unobservable (omitted), the estimate of β_1 can be expressed as:

$$b_1 = \left(X_1'X_1\right)^{-1} X_1'Y$$

$$b_1 = \left(X_1'X_1\right)^{-1} X_1'\left(\beta_1 X_1 + \beta_2 X_2 + \varepsilon\right)$$

$$b_1 = \left(X_1'X_1\right)^{-1}\left(X_1'X_1\right)\beta_1 + \left(X_1'X_1\right)^{-1}\left(X_1'X_2\right)\beta_2 + \underbrace{\left(X_1'X_1\right)^{-1} X_1'\varepsilon}_{=0} \tag{3}$$

From Equation 3, b_1 is a true estimate of β_1 only if $\left(X_1'X_1\right)^{-1}\left(X_1'X_2\right)\beta_2 = 0$, for which either $\left(X_1'X_1\right)^{-1}\left(X_1'X_2\right) = 0$, or $\beta_2 = 0$, or both. Autor et al. (2019) and Xu and Gong (2017) show that U.S. firms experience a decline in patenting rates and R&D expenditures in response to increase in import competition, which implies that $\beta_2 < 0$. On the other hand, using similar mcirodata as this paper, Magyari (2017) shows that increased Chinese imports in U.S. input markets acted as a favorable cost shock to U.S. manufacturing firms, which implies that offshoring and import competition are positively correlated, i.e. $\beta_2 > 0$ and OLS coefficients are liklely to be downward biased consistent with Bernard et al. (2018):

$$E[b_1|X_1] = \beta_1 + \underbrace{\underbrace{\left(X_1'X_1\right)^{-1} X_1'X_2}_{>0}\,\underbrace{\beta_2}_{<0}}_{<0}$$

4 Reduced-Form Empirical Strategy and Results

My primary reduced-form empirical strategy is to relate changes in within firm innovation and employment outcomes to exogenous changes in offshoring, after after detailed industry and time fixed fixed effects and other time varying technology augmenting firm-level scale factors like size, sales, capital stock and labor productivity. To that end, I estimate the following regression for firm i, in NAICS six industry j[22], at time t:

$$\ln \mathrm{Y}_{ijt} = \beta_{\mathrm{O}} \ln \mathrm{Off}_{it}^{\mathrm{type}} + \beta_{\mathrm{X}} \ln \mathrm{X}_{it} + \theta_i + \gamma_j + \delta_t + \varepsilon_{it} \tag{4}$$

where $type \in (Total, Related, Total\ China, Related\ China)$. The outcome variables I consider in Equation 4, are firm R&D expenditures, average earnings of R&D workers, production workers and non-production workers. Following Boler et al. (2015), the vector X_{ijt} contains firm-level contemporaneous controls like size, sales, capital stock and labor productivity (all in logs), as potential determinants of innovation and employment outcomes. Using firm (θ_i), industry (γ_j) and time (δ_t) fixed effects, the coefficient of interest in Equation 4 is β_{O} which measures the marginal effect of within-firm offshoring on the outcome variable of interest. With a log-log specification, the marginal effect can be interpreted as a 1 percentage point change in offshoring resulting in a β_{O} percentage point change in the outcome variable of interest. Note, that I instrument for the different types of offshoring using a shift-share design according to subsection 3.2.

A limitation with the R&D data stems from the survey design implemented in the SIRD/BRDIS. In addition to being biased towards R&D performing firms, it surveys firms intermittently depending on their R&D involvement. For example, a firm might be surveyed consecutively from 2003 to 2008 before being dropped from the survey in 2009, to be introduced back again in 2011. This precludes the interpretation of the marginal effect as an annual change, since the interval might be more than 1 year for some firms.

[22] The Longitudinal Business Database provides a time varying detailed six digit NAICS industry for each establishment which depends on the primary activity performed within the establishment. In order to construct a consistent, time invariant NAICS code for the firm, I choose the industry with the modal employment over the lifecyle of the firm.

More specifically, 61% of the firm-year sample is continuous in nature, while 39% of firm-year observations have breaks of atleast two years or more. Hence the marginal effect of β_{O} in Equation 4 is not a result of an uniform annual change precisely. I use the entire sample in estimating Equation 4 to provide a sense of the endogenity (bias), instrumentation and overall relationship between plausibly exogenous measures of offshoring and R&D expenditures.

In order to faciliate the interpretation of the marginal effect, I proceed by estimating Equation 4 in annualized first differences. I annualize the changes in R&D, offshoring and controls, to interpret β_{O} as the marginal effect of the growth of offshoring (annual) on the growth of R&D expenditures. Estimating the model in first differences also controls for unobserved firm heterogeneity, as firm level fixed effects cancel out. I include growth of scale factors (like sales, capital and labor productivity) to control for unobserved firm level factors affecting the growth of R&D. Additionally, instead of separate industry and year fixed effects, I use six-digit NAICS $\times$ year fixed effects to control for any potential time varying industry level shocks like import competition or technology/demand shocks affecting R&D growth of firms. More specifically I estimate,

$$\Delta \ln \mathrm{Y}_{ijt} = \beta_{\mathrm{O}}\ \Delta \ln \mathrm{Off}_{it}^{\mathrm{type}} + \beta_{\mathrm{X}}\ \Delta \ln \mathrm{X}_{it} + \Delta\gamma_{jt} + \Delta\varepsilon_{it} \tag{5}$$

where $\Delta \ln \mathrm{Y}_{ijt} = \ln \mathrm{Y}_{ijt} - \ln \mathrm{Y}_{ijt-1}$. For cases, where firm-year observations are non-consecutive, I annualize the difference in dependent and independent variables by dividing by the number of years. Finally, in response to concerns of differential trends across firms, as an exhaustive specification, I also estimate a model with firm-specific trends, sometimes referred to as a correlated random trend model.[23]

4.1 R&D Expenditure

In order to relate changes in R&D outcomes to offshoring, I first estimate Equation 4 using log of R&D expenditures as the outcome variable. Equation 4 is estimated on the full sample of firm-year observations, including inconsistent time intervals within firms. The purpose of this exercise is to show the overall direction of the effect of offshoring, and the complementary effect of sourcing more similar goods on R&D outcomes. The β_{O} coefficient is expected to be positive indicating that increase in within-firm offshoring is associated with increase in firm level innovation. However, due to discontinuity in the panel structure, I refrain from interpreting β_{O} as the marginal effect of annual change in offshoring on R&D.

Few important patterns in the data stands out. First, after controlling for firm-level scale factors like capital stock, sales and labor productivity, along with year and industry fixed effects, Table 6 column (1) shows that offshoring has a positive effect on R&D expenditures across firms. However, identifying within-firm variation in column (2), the effect is significantly attenuated and is no longer statistically distinguishable from zero. Second, as argued in subsection 3.2, offshoring is endogenous and OLS coefficients in addition to simultaneity bias, could be potentially biased upward(downward) in the presence of positive(negative) demand/productivity shocks. In order to account for such competing channels, I follow Hummels et al. (2014) by using a shift-share design instrument where I apportion changes in world export supply to firms based on their pre-sample destination, product specific import shares as weights according to Equation 1. Table 6 column (3) shows that across firms, the IV estimate of offshoring is almost twice in magnitude compared to its OLS counterpart. Further, on including firm fixed effects and thereby identifying within-firm variation in offshoring, the marginal effect in column (4) is robust and significant at the 1% level.[24] Comparison of columns (2) and (4) confirms evidence of downward bias in OLS, similar to

[23]Hence the FD model of Equation 4 with firm specific trends leads to: $\Delta \ln \mathrm{Y}_{ijt} = \beta_{\mathrm{O}}\ \Delta \ln \mathrm{Off}_{it}^{\mathrm{type}} + \beta_{\mathrm{X}}\ \Delta \ln \mathrm{X}_{it} + \theta_i + \Delta\delta_t + \Delta\varepsilon_{it}$

[24]Comparing columns (2) and (4), I find that the IV effect is roughly ten times larger in magnitude, similar to Hummels

Table 6: **Baseline Results - Offshoring & R&D Expenditure (Full Sample)**

	Log R&D Expenditure			
	(1)	(2)	(3)	(4)
Log Offshoring (4 digit)	0.129***	0.011	0.201***	0.104***
	(0.010)	(0.011)	(0.019)	(0.030)
Estimation	OLS	OLS	IV	IV
Firm Fixed Effect	X	✓	X	✓
Industry Fixed Effect	✓	✓	✓	✓
Year Fixed Effect	✓	✓	✓	✓
Firm Controls	✓	✓	✓	✓
WES			fpc	fpc
First Stage F-stat			680.1	25.70
R^2	0.465	0.882	0.463	0.881
Observations	81000	67000	81000	67000

Notes: The dependent variable is log **R&D expenditures** from 1997 to 2015. The independent variables are log "narrow" offshoring restricted at the 4 digit level as that of firm shipments. *WES* denotes World Export Supply defined in subsection 3.2. Firm controls include log of capital stock, log of shipments and log of labor productivity. In columns (3) and (4), instrument weights are firm-product-country(fpc). Industry refers to NAICS six digit classification. All regressions are weighted by SIRD/BRDIS sample weights and include a constant. All standard errors are clustered at the firm level.
*** Significant at the 1% level; ** Significant at the 5% level; * Significant at the 10% level.

Bernard et al. (2018).

In order to interpret the marginal effect in terms of an annual percentage change, I proceed by estimating Equation 5, where all variables are transformed to annualized first differences. Additionally, the exclusion restriction of the instrument requires changes in world export supply to affect firm-level R&D expenditure only through the offshoring channel. In order to prevent industry level shocks as argued by Bloom, Draca and Van Reenen (2015) and Autor et al. (2019) to affect R&D outcomes, thereby biasing the marginal effect of offshoring, I include detailed six-digit NAICS by year fixed effects to absorb any time varying industry shocks like import competition and export access, in addition to the existing set of contemporaneous controls.[25]

Table 7 presents estimates from the preferred specification in Equation 5. Comparison of columns (1) and (2) once again shows evidence of downwards bias in OLS as in Table 6. I interpret the marginal effect in column (2) as the baseline effect, which translates to a slightly more than 0.7% average increase in R&D, in response to a 10% increase in total offshoring. While I delay precise statements about the economic significance of the marginal effects until later, two other patterns stand out in the data. In columns (3) through (6), I decompose the total effect of offshoring into related party and arms-length trade transactions. While arms-length trade transactions constitute majority of offshoring for most firms,[26] there is substantial evidence in the literature regarding offshoring, productivity and ownership of foreign partners (Kohler and Smolka (2009), Tomiura (2007)). Intra-firm or related party imports, aside of being correlated with firm size, allows firms an opportunity to reduce import uncertainly and transfer design and production knowledge to foreign firms to better customize inputs required for production. Column (4) shows that related-party imports, although less in absolute value, does have a significant effect on R&D expenditures, while arms-length offshoring closely mirrors the effect of total offshoring and is significant at the 1% level.

et al. (2014).

[25]Results remain nearly unchanged if I restrict the sample to consecutive year observations within firms. Results available upon request.

[26]Related party or Intra-firm imports are measured from the Related Party Trade Statistics published by the U.S. Census Bureau. Intra-firm trade is defined as trade between two parties where one party holds at least a 6 percent ownership stake in the other.

Table 7: Offshoring & R&D Expenditure: First Differences

	Log R&D Expenditure					
	(1)	(2)	(3)	(4)	(5)	(6)
Total Offshoring (4 digit)	0.008	**0.073*****				
	(0.007)	**(0.026)**				
Related Offshoring (4 digit)			-0.005	0.039**		
			(0.006)	(0.017)		
Arms-Length Offshoring (4 digit)					0.007	0.075***
					(0.007)	(0.028)
Estimation	OLS	IV	OLS	IV	OLS	IV
Firm Fixed Effect	X	X	X	X	X	X
Industry × Year Fixed Effect	✓	✓	✓	✓	✓	✓
Firm Controls	✓	✓	✓	✓	✓	✓
WES		fpc		fpc-rel		fpc-arms
R^2	0.305	0.302	0.305	0.304	0.305	0.302
Observations	51000	51000	51000	51000	51000	51000

Notes: The dependent variable is log **R&D expenditures** in first differences from 1997 to 2015. The independent variables are log "narrow" offshoring (in first differences) restricted at the four digit level as that of firm shipments. *WES* denotes World Export Supply defined in subsection 3.2. In columns (4) and (6), weights are restricted to pre-sample related-party and arms-length product country specific imports respectively. Firm controls include log of capital stock, log of shipments and log of labor productivity. Industry refers to NAICS six digit classification. All regressions are weighted by SIRD/BRDIS sample weights and include a constant. All standard errors are clustered at the firm level.
******* Significant at the 1% level; ****** Significant at the 5% level; ***** Significant at the 10% level.

The idea of "narrow offshoring" defined by Feenstra and Hanson (1999) and later applied to firms by Hummels et al. (2014) argued that closer the inputs are to the final outputs, it is more likely that labor within the firm could have produced those inputs. In this context, it implies that more similar inputs, allow firms the opportunity to substitute domestic labor with foreign production and redirect domestic resources to innovation. I formally test this by showing that, closer the firm imports are to their domestic shipments, the stronger is the effect of offshoring on domestic innovation. Table 8 shows that the effect of offshoring on R&D monotonically increases with the degree of similarity between firm imports and shipments. The final and most conservative measure of offshoring which requires firm imports to be in the exact six-digit NAICS industry as that of its shipments, closely reflects the "produced-good" imports used by Bernard et al. (2018).

4.2 Earnings

A significant fraction of firm level R&D expenditure occurs in the form of salaries and wages to high skilled engineers and R&D personnel. If U.S. firms focus on design and R&D, rather than on the physical transformation of goods, they are likely to hire skilled workers and pay them high wages in response to increased input sourcing from abroad. However, it is not straightforward to sort out the effect on manufacturing earnings empirically. If production of "high-quality" intermediates is retained at home, then one can expect selection of highly productive manufacturing workers retaining their jobs and thereby earning more. Additionally, as suggested by Bloom, Draca and Van Reenen (2015), manufacturing plants react to import competition by undertaking organizational innovations, that may raise productivity and wages.

As with R&D expenditures, I start by using the full sample of firm-year observations to highlight key trends in the data Table 9 shows that while offshoring has a positive and significant effect on within-firm average R&D earnings, there is virtually no effect on average earnings of production workers. The positive effect on R&D earnings does support the idea of a within-firm organizational change in response to

Table 8: **Offshoring & R&D Expenditure: First Differences**

	Log R&D Expenditure		
	(1)	(2)	(3)
Log Offshoring (4 digit)	**0.073***** **(0.026)**		
Log Offshoring (5 digit)		0.082*** (0.029)	
Log Offshoring (6 digit)			0.108*** (0.038)
Estimation	IV	IV	IV
Firm Fixed Effect	X	X	X
Industry × Year Fixed Effect	✓	✓	✓
Firm Controls	✓	✓	✓
WES	fpc	fpc	fpc
R^2	0.302	0.301	0.298
Observations	51000	51000	51000

Notes: The dependent variable is log **R&D expenditures** from 1997 to 2015. The independent variables are log "narrow" offshoring restricted at the 4-6 digit levels as that of firm shipments. *WES* denotes World Export Supply defined in subsection 3.2. Firm controls include log of capital stock, log of shipments and log of labor productivity. Industry refers to NAICS six digit classification. All regressions are weighted by SIRD/BRDIS sample weights and include a constant. All standard errors are clustered at the firm level.
*** Significant at the 1% level; ** Significant at the 5% level; * Significant at the 10% level.

Table 9: **Offshoring & Average Earnings (Full Sample)**

	Log R&D Earnings		Log Production Earnings	
	(1)	(2)	(3)	(4)
Log Offshoring (4 digit)	0.063*** (0.008)	0.256*** (0.020)	-0.001 (0.001)	-0.001 (0.001)
Estimation	OLS	IV	OLS	IV
Firm Fixed Effect	✓	✓	✓	✓
Industry Fixed Effect	✓	✓	✓	✓
Year Fixed Effect	✓	✓	✓	✓
Firm Controls	✓	✓	✓	✓
WES weight		fpc		fpc
First Stage F-Stat		55.74		36.83
R^2	0.795	0.787	0.873	0.873
Observations	42000	42000	57500	57500

Notes: The dependent variables are log **R&D earnings** (columns (1) and (2)) and log **Production earnings** (columns (3) and (4)) from 1997 to 2013. The independent variable are log "narrow" offshoring restricted at the 4 digit level as that of firm shipments. *WES* denotes World Export Supply defined in subsection 3.2. Firm controls include log of capital stock, log of shipments and log of labor productivity. Industry refers to NAICS six digit classification. All regressions are weighted by SIRD/BRDIS sample weights and include a constant. All standard errors are clustered at the firm level.
*** Significant at the 1% level; ** Significant at the 5% level; * Significant at the 10% level.

offshoring, by focusing on high-valued added activities like design and R&D and thereby paying premium wages to high-skilled workers. On the other hand, the insignificant effect on production earnings echoes the findings in Autor, Dorn and Hanson (2013), who find manufacturing earnings to be unresponsive to local trade shocks. This finding also lines up with evidence in Autor et al. (2019), that U.S. firms did not respond to increase in Chinese import competition by patenting and R&D, which could possibly hike manufacturing wages.

I proceed by estimating the effect of offshoring on R&D and production earnings, in first differences. Table 10 column (2) shows that a 10% increase in offshoring on average is likely to cause a 0.94% increase in R&D earnings, thereby highlighting a rise in skill premium associated with sourcing of foreign interme-

Table 10: Offshoring & Average Earnings: First Differences

	Log R&D Earnings		Log Production Earnings	
	(1)	(2)	(3)	(4)
Log Offshoring (4 digit)	0.019***	**0.094*** **	-0.000	**-0.000**
	(0.007)	**(0.021)**	(0.000)	**(0.001)**
Estimation	OLS	IV	OLS	IV
Firm Fixed Effect	X	X	X	X
Industry × Year Fixed Effect	✓	✓	✓	✓
Firm Controls	✓	✓	✓	✓
WES weight		fpc		fpc
R^2	0.419	0.415	0.211	0.211
Observations	31500	31500	31500	31500

Notes: The dependent variables are log **R&D earnings** (columns (1) and (2)) and log **Production earnings** (columns (3) and (4)) from 1997 to 2013. The independent variable are log "narrow" offshoring restricted at the 4 digit level as that of firm shipments. *WES* denotes World Export Supply defined in subsection 3.2. Firm controls include log of capital stock, log of shipments and log of labor productivity. Industry refers to NAICS six digit classification. All regressions are weighted by SIRD/BRDIS sample weights and include a constant. All standard errors are clustered at the firm level.
******* Significant at the 1% level; ****** Significant at the 5% level; ***** Significant at the 10% level.

diates. Similar to R&D expenditure, I test if more closely related imports have a stronger impact on R&D earnings. In Table 11, I find that import of inputs that are in the exact same six digit NAICS industry as that of firms' shipment, have the strongest effect on R&D earnings, implying greater substitutability of production and reallocation towards innovative activities.

Table 11: Offshoring & Average Earnings: First Differences

	Log R&D Earnings		
	(1)	(2)	(3)
Log Offshoring (4 digit)	**0.094*** **		
	(0.021)		
Log Offshoring (5 digit)		0.104***	
		(0.024)	
Log Offshoring (6 digit)			0.138***
			(0.032)
Estimation	IV	IV	IV
Firm Fixed Effect	X	X	X
Industry × Year Fixed Effect	✓	✓	✓
Firm Controls	✓	✓	✓
WES	fpc	fpc	fpc
R^2	0.415	0.412	0.404
Observations	31500	31500	31500

Notes: The dependent variables are log **R&D earnings** (columns (1) and (2)) and log **Production earnings** (columns (3) and (4)) from 1997 to 2013. The independent variable are log "narrow" offshoring restricted at the 4 digit level as that of firm shipments. *WES* denotes World Export Supply defined in subsection 3.2. Firm controls include log of capital stock, log of shipments and log of labor productivity. Industry refers to NAICS six digit classification. All regressions are weighted by SIRD/BRDIS sample weights and include a constant. All standard errors are clustered at the firm level.
******* Significant at the 1% level; ****** Significant at the 5% level; ***** Significant at the 10% level.

In order to providence economic significance of estimates presented in Table 7 and Table 10, I consider the full distribution of changes in offshoring that occur in our sample, and the corresponding change in R&D expenditure and earnings. To that end, I follow Hummels et al. (2014) by categorizing firm-years into bins on the basis of annual changes in offshoring for that firm. I, then report for each bin, the share of firm-year observations and the median R&D expenditures (R&D earnings) experienced by workers as predicted using coefficient estimates from Table 7(Table 10). Consider the first row of Table 12. This corresponds to firm-years where annual change in offshoring is atleast 50%, which represent 21.6% of the sample used for estimation. Using the estimates from Table 7, I predict that firms will experience a median R&D expenditure annual increase of 23.66%. Similarly, for firms that experienced more than 50% increase in annual offshoring, will experience a median R&D earnings annual increase of 22.8%.[27]

Table 12: Economic Effect of Offshoring on R&D Outcomes

Bin	Sample Share (firm-year)	Median Predicted Change
	Panel A: R&D Expenditures (1997-2015)	
% Increase in Offshoring (Annual)		
$> 50\%$	21.6%	23.66%
$25 - 49\%$	7%	2.55%
$5 - 24\%$	11.1%	0.98%
% Decrease in Offshoring (Annual)		
$> 50\%$	9%	-6.28%
$25 - 49\%$	7.4%	2.62%
$5 - 24\%$	11.6%	-1%
	Panel A: R&D Earnings (1997-2013)	
% Increase in Offshoring (Annual)		
$> 50\%$	22.8%	22.38%
$25 - 49\%$	8.1%	3.31%
$5 - 24\%$	12.4%	1.29%
% Decrease in Offshoring (Annual)		
$> 50\%$	10.2%	-7.97%
$25 - 49\%$	8.4%	-3.77%
$5 - 24\%$	11.9%	-1.31%

From these patterns in the data, it is clear that a significant fraction of the R&D increase is accounted for by firms experiencing more than 50% increase in annual offshoring. However it is striking that such firm-years account for more than one-fifth of the sample, while more than 38% of the sample experience greater than 5% increase in annual offshoring. Note that in all regression specifications, I control for contemporaneous size, labor productivity and capital stock to account for the presence of large firms who engage in both R&D and offshoring.

Finally, I also estimate an exhaustive model with firm-specific random trends, after controlling for contemporaneous controls as Table 7, along with industry and year fixed effects. For comparison, I report estimates from the preferred specification in Table 7 and Table 10. In column (2), I find that the firm specific trend absorbs significant variation in R&D expenditures due to offshoring. The marginal effect declines by more than 50% and is only marginally significant at the 10% level. However, for average R&D earnings, including a firm-specific trend increases the marginal effect of offshoring. According to column (4), a 10% increase in offshoring is associated with an almost 1.2% increase in average R&D earnings.[28]

These results point to a skill biased, productivity augmenting role of offshoring, in turn explaining why a number of studies, in the past have found find large firm-level productivity gains associated with input trade liberalization (Boler, Moxnes and Ulltveit-Moe (2015)). It also highlights a global phenomenon

[27] For the most restrictive definition of six digit offshoring, I use estimates from column (3) in Table 8. Results are presented in Table 20. For firms experiencing more than 50% annual increase in six-digit offshoring, the median R&D increases by 43.72%.

[28] Controlling for the firm-specific trend has no effect on production earnings.

Table 13: **Offshoring & R&D Outcomes: First Differences (Robustness)**

	Log R&D Expenditure		Log R&D Earnings	
	(1)	(2)	(3)	(4)
Log Offshoring (4 digit)	**0.073***** **(0.026)**	0.039* (0.022)	**0.094***** **(0.021)**	0.118*** (0.023)
Estimation	IV	IV	IV	IV
Firm Fixed Effect	X	✓	X	✓
Year Fixed Effect	✓	✓	✓	✓
Firm Controls	✓	✓	✓	✓
WES weight	fpc	fpc	fpc	fpc
R^2	0.302	0.289	0.415	0.305
Observations	51000	49500	31500	29500

Notes: The dependent variables are log **R&D earnings** (columns (1) and (2)) and log **Production earnings** (columns (3) and (4)) from 1997 to 2013. The independent variable are log "narrow" offshoring restricted at the 4 digit level as that of firm shipments. *WES* denotes World Export Supply defined in subsection 3.2. Firm controls include log of capital stock, log of shipments and log of labor productivity. Industry refers to NAICS six digit classification. All regressions are weighted by SIRD/BRDIS sample weights and include a constant. All standard errors are clustered at the firm level.
******* Significant at the 1% level; ****** Significant at the 5% level; ***** Significant at the 10% level.

where firms in developed countries like the U.S., instead of focusing on physical transformation of goods, concentrate on areas of core competence within the headquarter country. For such firms, stages of core competencies often involve employment of skilled labor, high-tech manufacturing and innovation. Having robustly established the effect of offshoring on R&D outcomes, I then, develop and estimate a structual dynamic model of R&D to highlight, the channels through which this complementary effect of offshoring and R&D affect future firm performance.

5 Structural Dynamic Model and Estimation

Motivated by the reduced form findings in Section 4 on the link between offshoring and R&D, I build and estimate a structural dynamic model of R&D, following Aw, Roberts and Xu (2011), with R&D choice being the only dynamic decision facing firms, in their effort to augment future productivity.[29] Following Halpern, Koren and Szeidl (2015) and Zhang (2017), I introduce foreign intermediates in the production function, to analyze how a reduction in the relative cost of imported intermediates, leads to increased R&D participation, thereby endogenously affecting future firm performance. Imported intermediates enable firms to lower marginal costs of production and gain from potentially lower quality-adjusted price, access to a larger variety of imported inputs or a combination of both. Increase in firm profitability, accounted for by static gains from offshoring, in turn allows firms to engage in R&D, by overcoming the associated fixed and sunk costs. These investments have feedback effects that can potentially alter the path of future performance for the firm.

To the best of my knowledge, this is the first paper to link static gains from importing intermediates to dynamic gains from R&D, by explicitly estimating the decision rule of R&D and quantifying the complementary effect of offshoring and R&D on firm performance. The timing of the model is key to understanding the different components of firm decision-making. First, at the beginning of each period, each firm observes its state variables - capital stock, productivity and last period R&D status, following which it maximizes profit by optimally choosing labor, domestic and foreign inputs. Second, based on its lagged R&D status, it observes sunk and fixed cost of R&D, which are i.i.d. drawn from two different distributions. Based on their cost draws, firms decide whether to engage in R&D in the current period and pay fixed costs if they performed R&D last year, or incur sunk costs if they did not engage in R&D

[29] Aw, Roberts and Xu (2011) considered both R&D and export as dynamic decisions.

last year.

5.1 Static Decisions

Firms' production function is given by:

$$Y_{it} = L_{it}^{\beta_\ell} K_{it}^{\beta_k} M_{it}^{\beta_m} e^{\tilde{\omega}_{it}} e^{u_{it}} \tag{6}$$

where L_{it} is employment, K_{it} is capital stock, V_{it} is the quantity of an intermediate bundle and $\tilde{\omega}_{it}$ is a Hicks neutral productivity term. Following Halpern, Koren and Szeidl (2015), the composite intermediate bundle is assembled from a combination of domestic and foreign variety in the form of:

$$M_{it} = \left[(A_i M_{ift})^{\frac{\theta-1}{\theta}} + M_{idt}^{\frac{\theta-1}{\theta}} \right]^{\frac{\theta}{\theta-1}} \tag{7}$$

where M_{itd} and M_{itf} are domestic and foreign varieties respectively and θ is the elasticity of substitution between domestic and imported inputs. A_i can be interpreted as the price-adjusted quality advantage of the foreign input. The nested CES structure characterizes the static effect of importing through the *quality effect*(A) and the *variety effect*(θ). With the quality coefficient normalized to one, I do not restrict $A > 1$, since foreign intermediates sourced by U.S. firms can have potentially lower quality than domestic goods. The higher the elasticity of substitution(θ), implying higher substitutability, lower is the input variety effect.

Demand is assumed to be Dixit-Stiglitz type:

$$Q_{it}^D = \Phi_t P_{it}^{\eta}$$

where η is demand elasticity, P_{it} is firm i's price and Φ_t is the time variant demand shifter. Combining all importers and non-importers, in the spirit of Zhang (2017), the production function of the firm in Equation 6 can be written as:

$$Y_{it} = L_{it}^{\beta_\ell} K_{it}^{\beta_k} \left(\left[(A_i M_{ift})^{\frac{\theta-1}{\theta}} + M_{idt}^{\frac{\theta-1}{\theta}} \right]^{\frac{\theta}{\theta-1}} \right)^{\beta_m} e^{\tilde{\omega}_{it}} e^{u_{it}} \tag{8}$$

After log-linearizing and denoting $x = \ln x$, Equation 8 and can be expressed as:

$$y_{it} = \beta_\ell l_{it} + \beta_m \frac{\theta}{\theta - 1} \ln \left[(A_i M_{ift})^{\frac{\theta-1}{\theta}} + M_{idt}^{\frac{\theta-1}{\theta}} \right] + \beta_k k_{it} + \omega_{it} + \xi_{it} \tag{9}$$

Using firm level data on capital, labor and materials, I estimate the production parameters. Firm-level productivity ω_{it} is observable to the firm (unobservable to researchers), while ξ_{it} is unobserved iid measurement error. It is a well known problem that firm's choice of production inputs due to correlation with ω_{it}, is subject to problems of endogeneity. Using insights from Olley and Pakes (1996) and Levinsohn and Petrin (2003), I use data on firms' investment (levels) to obtain information on productivity ω_{it} and recover an unbiased estimate of firm productivity under the usual monotonicity assumption. More specifically, firm investment is modeled as a function of capital stock, productivity and future R&D status, $i_{it} = i_t(\omega_{it}, k_{jt}, rd_{it+1})$ which can be used to recover $\omega_{it} = \omega_t(i_{jt}, k_{jt}, rd_{it+1})$.

$$y_{it} = \beta_\ell l_{it} + \beta_m \frac{\theta}{\theta - 1} \ln \left[(A_i M_{ift})^{\frac{\theta-1}{\theta}} + M_{idt}^{\frac{\theta-1}{\theta}} \right] + \phi(i_{it}, k_{it}, rd_{it+1}) + \xi_{it} \tag{10}$$

where $\phi(i_{it}, k_{it}, rd_{it+1})$ estimates the combined effect of productivity and capital on production. Parameterizing $\phi(i_{it}, k_{it}, rd_{it+1})$ as a cubic function in capital, investment and future R&D status, Equation 10 can be estimated using any non-linear estimator in the first step, to estimate the static input shares β_ℓ (labor), β_m (materials), the quality parameter $\hat{A}$, the variety parameter $\hat{\theta}$ and the $\phi(\cdot)$ function. Firm productivity is estimated in the second stage once I specify a structure for the evolution of endogenous productivity.

<u>*Transition Process:*</u> Following Aw, Roberts and Xu (2011), Doraszelski and Jaumandreu (2013) among others, I assume firm performance to evolve over time following a controlled first-order Markov process, that depends on whether a firm conducts R&D, as well as a random shock:

$$\omega_{it} = \alpha_0 + \alpha_1\omega_{it-1} + \alpha_2 rd_{it-1} + \varepsilon_{it} \qquad \text{with iid } \varepsilon_{it} \sim N(0, \sigma_\varepsilon^2) \tag{11}$$

where $rd_{it-1} = 1$ if the firm engaged in R&D the period before, and 0 otherwise.[30] The α_2 coefficient embodies the idea that R&D has a dynamic effect on productivity, which in turn affects future productivity through α_1.

Given $\hat{\phi}(\cdot)$ (from Equation 10), I can express $\omega_{it} = \hat{\phi}_{it} - \beta_k k_{it}$ where $\hat{\phi}_{it-1}$ is the estimate of ϕ_{it-1}. Substituting ω_{it} in Equation 11 and replacing ω_{it-1} with $\omega_{it-1} = \hat{\phi}_{it-1} - \beta_k k_{it-1}$ yields the estimating equation:

$$\phi_{it} = \alpha_0 + \beta_k k_{it} + \alpha_1\left(\hat{\phi}_{it-1} - \beta_k k_{it-1}\right) + \alpha_2 rd_{it-1} + \varepsilon_{it} \tag{12}$$

Now, I can consistently estimate β_k and given the estimate of β_k, I can construct productivity estimates for each firm-year observation:

$$\hat{\omega}_{it} = \hat{\phi}_{it} - \hat{\beta}_k k_{it} \tag{13}$$

This static model similar to Zhang (2017) which builds on Olley and Pakes (1996), has the added benefit of estimating the *quality* parameter $\hat{A}$ and the *variety* parameter $\hat{\theta}$ in addition to the standard elasticities of capital, labor, materials and the pseudo productivity sample. Both the quality and variety parameters are of special interest in my analysis because of two reasons. On one hand, it sheds light on the margins along which U.S. firms benefit from importing (either via import of higher price adjusted quality intermediates or via increased access to higher variety of intermediates). On the other hand, these structural parameters help in counterfactual analysis by altering static gains from trade and analyzing how it affects other dynamic firm-level decisions.

Finally, I follow Zhang (2017) to express the firm's revenue function in terms of productivity, capital stock, current importing status and a time-varying effect γ_t which captures aggregate demand shocks. More specifically:

$$\ln \text{Y}_{it} = \gamma_t + r_k \ln \text{K}_{it} + r_\omega\omega_{it} + r_d d_{it} \tag{14}$$

where $r_k = \frac{\beta_k}{\frac{\eta}{(1+\eta)} - (\beta_\ell + \beta_m)}$, $r_d = \left[\frac{\alpha_m/(\theta-1)}{\frac{\eta}{(1+\eta)} - (\beta_\ell+\beta_m)}\right] \ln\left[1 + \left(A\frac{P_{dt}}{P_{ft}}\right)^{\theta-1}\right]$ and $r_\omega = \frac{(\beta_\ell+\beta_m)}{\frac{\eta}{(1+\eta)} - (\beta_\ell+\beta_m)}$. r_d is the impact of importing on current revenue which depends on both the quality parameter (A) and variety parameter(θ). Accordingly the total variable cost and profit is a fixed share of revenue:[31]

$$C_{it} = \frac{1+\eta}{\eta}(\beta_\ell + \beta_m) Y_{it} \tag{15}$$

$$\pi_{it} = \left[1 - \frac{1+\eta}{\eta}(\beta_\ell + \beta_m)\right] Y_{it} \tag{16}$$

[30] In a robustness check, I also include exports in the productivity evolution Markov process. Results remain qualitatively unchanged. Available upon request.

[31] Refer to Zhang (2017) for derivation and details.

To summarize, I outline a simple static model to estimate the firms' production function along with two key parameters - the import quality effect (A) and the variety effect (θ) which directly affect current revenue and hence firm value. Next, I estimate the productivity evolution process, thereby estimating the contribution of R&D choice on future productivity, which help in governing endogenous R&D choice of firms. The corresponding firm profits help in computing value functions of firms, which is discussed later.

5.2 Dynamic Decisions

In this section, I outline a dynamic structural model analyzing firm's endogenous decision to engage in R&D, in an attempt to boost firm performance. To that end, the model closely follows Aw, Roberts and Xu (2011) in modeling firms' R&D participation with random sunk and fixed costs of R&D. However unlike Aw, Roberts and Xu (2011), I do not model the dynamic export market performance of firms. R&D is the only dynamic component in the model, as shown in the productivity evolution process in Equation 11.

To analyze the dynamic decision of firms to engage in R&D, the main considerations are the dynamic gains from R&D, and fixed and sunk costs associated with undertaking R&D. Lagged R&D status is relevant since first-time R&D performers need to pay a sunk cost while, R&D continuing firms pay per period fixed costs. It is reasonable to think that sunk costs are higher than fixed costs and I assume that both costs are i.i.d. draws from two different distribution: Sunk costs ($C^s_{it} \sim F^s$) and fixed costs ($C^f_{it} \sim F^f$). A R&D performing firm needs to incur these costs to enjoy productivity gains, as denoted by the coefficient α_2 in Equation 11. The across-firm variation in these expenditures stem from differences in technological expertise, and I assume that the firm observes its expenditure before making its discrete decision to invest in R&D. The endogeneity in firm performance arises from the endogenous dynamic decision of firms to engage in R&D, which in a recursive formulation (Equation 11) also affects future performance.

The fixed and sunk costs will be treated as containing a stochastic component, which the firm observes prior to making its R&D decision in the current period, but is not observed by the researcher. This makes the firm's past R&D participation a state variable in the dynamic model. Hence the state vector for firm i in year t is $s_{it} = (\omega_{it}, k_{it}, rd_{it-1})$[32] and the firm's value function before it observes its fixed and sunk costs is:

$$V(s_{it}) = \int\int \left[\pi(s_{it}) + \max_{rd_{it}}\{V^1(s_{it}) - rd_{it-1}C^f_{it} - (1 - rd_{it-1})C^s_{it}, V^0(s_{it})\}\right] dF^s dF^f \tag{17}$$

where rd_{it-1} is a discrete 0/1 variable denoting whether the firm invested in R&D last period in $t-1$. If the firm engaged in R&D last period, then it is considered as a R&D continuing firm and pays fixed cost C^f_{it}. Otherwise it is considered as a R&D starter and accordingly pays R&D sunk cost C^s_{it}. In that sense, last period R&D status is dynamic not only though productivity (Equation 11) but also through the cost channel (Equation 17). To be more precise the choice specific value functions can be written as

$$V^1(s_{it}) = \delta E_t V(s_{it+1} \mid s_{it}, rd_{it} = 1) \tag{18}$$

$$V^0(s_{it}) = \delta E_t V(s_{it+1} \mid s_{it}, rd_{it} = 0) \tag{19}$$

where δ is the discount factor.

[32] Current capital equals current investment plus last period's capital stock after depreciation: $K_{it} = (1-\delta)K_{it-1} + I_{it}$.

5.3 Estimation Strategy and Results

In order to estimate the static parameters in the production function, I use data on firm level employment, capital, revenue,[33] domestic and imported material, and a 0/1 importing indicator.[34] A limitation with U.S. trade and production data is that, there is no information on the share of materials that is imported or produced/sourced domestically. I assume that imports of goods that are similar to the firm shipment goods atleast at the four digit level, comprise firm material imports, while the remainder of total materials is denoted as domestic material. Although not perfect, this definition allows me to quantify a lower bound on imported materials and also exploit substantial variation in imports across firms, both along the extensive and intensive margins.

The set of static parameters to be estimated include: β_ℓ, β_m, A and θ, of which A and θ are of major interest since they quantify static gains from importing due to the *quality* and *variety* effect respectively. I implement the static estimation in two ways. First, I pool firms across all years, and estimate the average production parameters similar to Halpern, Koren and Szeidl (2015). These estimates provide an overall sense of input elasticities, which can be compared to other microdata across countries to understand key dimensions along which they differ. Second, I estimate the model in two sub-samples, 1997-2006 and 2007-2015. The first period coincides with China's accession to the WTO and the expansion of U.S. firms' global intermediate sourcing strategies. Other than the financial crisis, the second period saw relative stability in trade costs, with increasing quality of exports especially from China.

Table 14: First Stage Estimates

Sample	Sub-Period 1 (1997-2006)	Sub-Period 2 (2007-2015)
β_ℓ	0.213	0.225
	(0.012)[1]	(0.009)
β_m	0.577	0.585
	(0.014)	(0.011)
A	0.544	0.955
	(0.18)	(0.186)
$\frac{\theta-1}{\theta}$	0.603	0.788
	(0.078)	(0.788)
θ	2.519	4.717

[1] Standard errors are reported in parentheses.

Table 14 reports the production parameters estimated from Equation 11. The estimated output elasticities, β_ℓ, β_m and β_k are all within reasonable range in terms of magnitude. Of particular interest, are the parameters A and θ which quantify the *quality* and *variety* gains from importing respectively. Two interesting findings stand out from Table 14. First, across both sub-samples, the quality parameter is less than one, significant in the first sub-period, while it is statistically indistinguishable from 1 in the second sub-period. It is important to note that A contains both the reality input quality effect and the price difference between domestic and foreign inputs. Hence, a statistically insignificant effect implies that the price differential accounts for the quality difference between domestic inputs and their foreign counterparts. This confirms Bernard et al. (2018), who find that offshorers import lower quality inputs, as reflected in their lower prices. However in the first sub-period, I find evidence of lower average quality even after adjusting for price differences. The *quality* effect also contrasts previous studies by (Halpern,

[33] I use level of firm shipments from the ASM/CM which is essentially not the same as total firm revenue.
[34] The importing indicator equal 1 if the value of 4 digit narrow offshoring is positive.

Koren and Szeidl, 2015) and Zhang (2017) who use data from developing countries to estimate a positive quality effect of foreign intermediates. Although less than one, the price adjusted quality doubles, when comparing across the two sub-periods.

Second, $\frac{\theta-1}{\theta} < 1$, confirming the existence of the variety effect, consistent with Goldberg et al. (2010). The estimated value of θ increases in the second sub-sample implying that the degree of substitutability between domestic and foreign inputs has increased over time. This is also consistent with the finding that the quality of U.S. imports has also increased, in line with Fan, Li and Yeaple (2015).[35]

Table 15 reports estimates from the productivity evolution process described in Equation 11.[36] Of particular interest, is the parameter α_2 which quantifies the dynamic effect of R&D choice on future productivity. $\alpha_2 > 0$ and statistically significant in both sub-periods, confirming the dynamic productivity effect of R&D, similar to Aw, Roberts and Xu (2011) and Boler, Moxnes and Ulltveit-Moe (2015). I estimate an overall demand elasticity ≈ 6.5 which is well within De Loecker's estimates ranging from 3 to 7.

Table 15: Second Stage Estimates

Sample	Sub-Period 1 (1997-2006)	Sub-Period 2 (2007-2015)
β_k	0.19	0.188
	(0.001)	(0.001)
	Panel A: Estimates of Markov process	
α_0	0.643	0.507
	(0.02)	(0.019)
α_1	0.553	0.643
	(0.014)	(0.013)
α_2	0.06***	0.024***
	(0.004)	(0.001)
σ_ω	0.068	0.056
	Panel B: demand elasticity	
η	-6.143	-8.23
	(0.001)	(0.001)

The static framework following Zhang (2017), differs from Halpern, Koren and Szeidl (2015) and Boler, Moxnes and Ulltveit-Moe (2015), in terms of proxying for intermediate imports. Instead of using levels of imported inputs, both Halpern, Koren and Szeidl (2015) and Boler, Moxnes and Ulltveit-Moe (2015) proxy for imports with a functional form, which takes into account the number of imported products at the firm-level. As a robustness mechanism, I use the number of distinct HS4 imported products, to obtain a similar static revenue effect of importing intermediates. Using the number of imported products (HS4) as a similar proxy for offshoring is reasonable, if it is correlated with the share of imported materials. In order to account for zero imported products, the alternate measure of offshoring is specified as $\ln(1 + n_{it})$ where n_{it} denotes the number of products imported by firm i at time t. To ensure that this alternate measure closely tracks the relationship between import share and the number of imported products, I plot the average import share of firms[37] as a function of n.

[35] On the full sample, I estimate the value of θ to be ≈ 3.5 which is close to Halpern, Koren and Szeidl (2015) estimate of 4.

[36] I also control for firms' export choice in the controlled first-order Markov process. Results are unchanged. Available upon request.

[37] The calculation of firm-level import share is not straightforward. The ASM and CM does not provide any information with respect to the total amount of imported intermediates. Additionally, the LFTTD does not specify if a particular product is used as an intermediate or final good. In order to proxy for imported intermediates, I use imports that qualify for the "narrow" measure of offshoring at the four digit level.

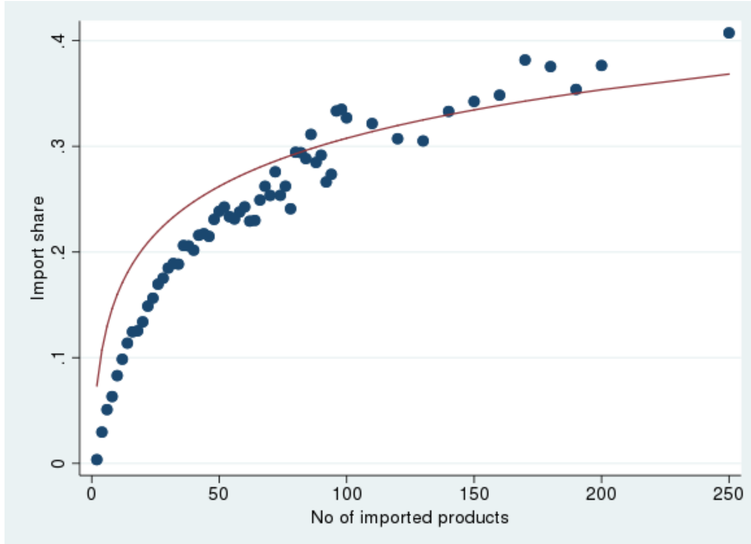


Figure 2: Import Share as a function of n_{it}

Armed with this alternate measure of offshoring, I estimate the firms' production function using a two-step method as in Boler, Moxnes and Ulltveit-Moe (2015).[38] Production function estimates presented in Table 16 are strikingly similar to those in Boler, Moxnes and Ulltveit-Moe (2015) using microdata from Norway.[39] In fact, the effect of imported products (≈ 10.06) is exactly the same in magnitude.

Table 16: Offshoring & Average R&D Earnings (Full Sample)

	One-step GMM	n exogenous	Continuous R&D
Capital(β_k^*)	0.736***	0.746***	0.730***
	(0.008)	(0.007)	(0.008)
No. imported products ($a\gamma^*$)	1.065***	0.943***	1.054***
	(0.039)	(0.031)	(0.039)
Productivity$_{t-1}$	0.312***	0.311***	0.320***
	(0.041)	(0.042)	(0.039)
Productivity$^2_{t-1}$	-0.010**	-0.010**	-0.009**
	(0.004)	(0.004)	(0.004)
R&D$_{t-1}$	0.042***	0.044***	0.004***
	(0.013)	(0.013)	(0.001)
Industry Fixed Effect	✓	✓	✓
Observations	40500	40500	40500

Notes: Standard errors are clustered at the firm level. R&D is a binary variable for columns (1) and (2) and log(1 + R&D expenditure) in column (3). Estimates of constant term omitted from table.
*** Significant at the 1% level; ** Significant at the 5% level; * Significant at the 10% level.

The distributions of sunk and fixed costs of R&D can be quantified from firms' dynamic decisions on whether or not to engage in R&D. After estimating the parameters of the production function along the demand elasticities (η), I derive the firms' profit function using Equation 16. Finally the sunk and fixed cost parameters can be identified from the conditional choice probabilities (CCP) of firms that begin or stop engaging in R&D. More specifically, a maximum likelihood estimate (MLE) is constructed for estimating the distribution of both sunk and fixed costs of R&D.

[38] Refer to Boler, Moxnes and Ulltveit-Moe (2015) for details about the model.
[39] The R&D effect is only for large firms. Small firm effect suppressed for disclosure requirements.

Given the state space $s_{it} = \{\omega_{it}, k_{it}, d_{it-1}\}$, the conditional probability of observing a firm with $rd_{it} = 1$:

$$\begin{aligned} L_{jt}^{1} &= \Pr\{rd_{it} = 1 \mid s_{it}\} \\ &= \Pr\left\{d_{it-1}C_{it}^{f} + (1 - d_{it-1})\,C_{it}^{s} \leq V^{1}(s_{it}) - V^{0}(s_{it}) \mid s_{it}\right\} \end{aligned} \tag{20}$$

where $V^1(s_{it})$ and $V^0(s_{it})$ are given by Equation 18 and Equation 19 respectively. Accordingly the probability of $rd_{it} = 0$ is $1 - L_{jt}^1$. Combining both, the probability of observing R&D status is:

$$L_{it} = d_{it}L_{it}^{1} + (1 - d_{it})(1 - L_{it}^{1}) \tag{21}$$

Extended to all firms, the probability of observing R&D status is:

$$L = \prod_{i=1:N} L_i = \prod_{i=1:N}\prod_{t=1:T} L_{it} \tag{22}$$

Summarizing the estimation procedure, I apply the nested fixed-point algorithm developed by Rust (1987) to solve for the value function $V^1(s_{it})$ and $V^0(s_{it})$. I discretize the state space into 25 grid points each for capital and productivity, and two values for lagged R&D discrete choice. Then I use the value function iteration in Equation 17 to solve for the value function at each point on the discretized state space. Finally, I apply a cubic spline to interpolate the firm value and payoff to R&D across the actual data space (there are XX firm-year observations in total).

6 Counterfactual Analysis

7 Conclusion

To be added.

References

Amiti, Mary, and Jozef Konings. 2007. "Trade Liberalization, Intermediate Inputs, and Productivity: Evidence from Indonesia." *American Economic Review*, 97(5): 1611–1638.

Autor, David, David Dorn, Gordon H Hanson, Gary Pisano, and Pian Shu. 2019. "Foreign Competition and Domestic Innovation: Evidence from U.S. Patents." *American Economic Review: Insights (forthcoming).*

Autor, David H., David Dorn, and Gordon H. Hanson. 2013. "The China Syndrome: Local Labor Market Effects of Import Competition in the United States." *American Economic Review*, 103(6): 2121–68.

Aw, Bee Yan, Mark J. Roberts, and Daniel Yi Xu. 2011. "R&D Investment, Exporting, and Productivity Dynamics." *American Economic Review*, 101(4): 1312–44.

Bernard, Andrew B, Teresa Fort, Valerie Smeets, and Frederic Warzynski. 2018. "Heterogeneous Globalization: Offshoring and Reorganization." *Working Paper.*

Bloom, Nicholas, Kyle Handley, Andre Kurmann, and Phillip Luck. 2019. "The Impact of Chinese Trade on U.S. Employment: The Good, The Bad, and The Debatable." *Working Paper.*

Bloom, Nicholas, Mirko Draca, and John Van Reenen. 2015. "Trade Induced Technical Change? The Impact of Chinese Imports on Innovation, IT and Productivity." *The Review of Economic Studies*, 83(1): 87–117.

Boehm, Christoph E., Aaron Flaaen, and Nitya Pandalai-Nayar. 2019. "Input Linkages and the Transmission of Shocks: Firm-Level Evidence from the 2011 TÅŊhoku Earthquake." *The Review of Economics and Statistics*, 101(1): 60–75.

Boler, Esther Ann, Andreas Moxnes, and Karen Helene Ulltveit-Moe. 2015. "R&D, International Sourcing, and the Joint Impact on Firm Performance." *American Economic Review*, 105(12): 3704–39.

Doraszelski, Ulrich, and Jordi Jaumandreu. 2013. "R&D and Productivity: Estimating Endogenous Productivity." *The Review of Economic Studies*, 80(4 (285)): 1338–1383.

Fan, Haichao, Yao Amber Li, and Stephen R. Yeaple. 2015. "Trade Liberalization, Quality, and Export Prices." *The Review of Economics and Statistics*, 97(5): 1033–1051.

Feenstra, Robert C., and Gordon H. Hanson. 1999. "The Impact of Outsourcing and High-Technology Capital on Wages: Estimates for the United States, 1979-1990." *The Quarterly Journal of Economics*, 114(3): 907–940.

Foster, Lucia, Cheryl Grim, and Nikolas Jason Zolas. 2016. "A Portrait of Firms that Invest in R&D." *US Census Bureau Center for Economic Studies Paper No. CES-WP-16-41.*

Ganapati, Sharat. 2017. "The Modern Wholesaler: Global Sourcing, Domestic Distribution, and Scale Economics." Yale mimeo.

Goldberg, Pinelopi Koujianou, Amit Kumar Khandelwal, Nina Pavcnik, and Petia Topalova. 2010. "Imported Intermediate Inputs and Domestic Product Growth: Evidence from India." *The Quarterly Journal of Economics*, 125(4): 1727–1767.

Halpern, Laszlo, Miklos Koren, and Adam Szeidl. 2015. "Imported Inputs and Productivity." *American Economic Review*, 105(12): 3660–3703.

Hummels, David, Rasmus Jorgensen, Jakob Munch, and Chong Xiang. 2014. "The Wage Effects of Offshoring: Evidence from Danish Matched Worker-Firm Data." *American Economic Review*, 104(6): 1597–1629.

Kasahara, Hiroyuki, and Joel Rodrigue. 2008. "Does the use of imported intermediates increase productivity? Plant-level evidence." *Journal of Development Economics*, 87(1): 106 – 118.

Kohler, Wilhelm K., and Marcel Smolka. 2009. "Global Sourcing Decisions and Firm Productivity: Evidence from Spain." *CESifo* Working Paper *No. 2903*.

Levinsohn, James, and Amil Petrin. 2003. "Estimating Production Functions Using Inputs to Control for Unobservables." *The Review of Economic Studies*, 70(2): 317–341.

Magyari, Ildiko. 2017. "Firm Reorganization, Chinese Imports, and US Manufacturing Employment." *Working Paper*.

Olley, G. Steven, and Ariel Pakes. 1996. "The Dynamics of Productivity in the Telecommunications Equipment Industry." *Econometrica*, 64(6): 1263–1297.

Tomiura, Eiichi. 2007. "Foreign outsourcing, exporting, and FDI: A productivity comparison at the firm level." *Journal of International Economics*, 72(1): 113–127.

Topalova, Petia, and Amit Khandelwal. 2011. "Trade Liberalization and Firm Productivity: The Case of India." *The Review of Economics and Statistics*, 93(3): 995–1009.

Xu, Rui, and Kaiji Gong. 2017. "Does Import Competition Induce R&D Reallocation? Evidence from the U.S." *International Monetary Fund, 2017, IMF Working Papers 17/253*.

Zhang, Hongsong. 2017. "Static and dynamic gains from costly importing of intermediate inputs: Evidence from Colombia." *European Economic Review*, 91: 118 – 145.

8 Appendix

FIRM LBD. The Longitudinal Business Database contains data on the universe of non-farm establishments in the U.S. over the period 1976-2015 conditional on having at least one employee on payroll over their life-cycle. Its underlying source is the Business Register (BR) which contains administrative records on U.S. businesses. At the establishment level, the LBD provides among other variables, data on annual employee count, payroll, detailed six digit NAICS industry, establishment age, location (zip, state and county codes) and most importantly an identifier connecting the establishment to a parent firm. Thus, it is straightforward to aggregate statistics to the firm level and the longitudinal nature of the dataset allows researchers to track establishments and firms over time.

LFTTD. A central piece of this analysis revolves around firm level import of intermediates. In order to do so, I exploit the Longitudinal Firm Trade Transactions Database (LFTTD), which contains import and export values (nominal US$) for the universe of trade transactions by detailed product code (HS 10 digit), source country (and destination) and firm identifier. Fortunately, the LBD and LFTTD share a common firm-level identifier which allows me to link the two databases conveniently over time.

PLANT CM. and ASM. A primary benefit of the LBD lies in its coverage both across establishments and over time. However, the LBD is limited in terms of its coverage of establishment/firm level variables. I augment establishment data on the count of production/non-production employees, levels of capital stock, investment, shipments and materials, by linking the Census of Manufacturers (CM for years ending in 7 and 2) and the Annual Survey of Manufacturers (ASM, for other years) to the LBD at the establishment level for years 1997-2015. Additionally, I link the ASM and CM to their respective product trailer files at the establishment level to obtain data on product shipments at an extremely disaggregated level (10 digit NAICS). On obtaining all the required information at the establishment level, I use the firm identifier in the LBD to aggregate the data at the firm level for further analysis.

SIRD. and BRDIS. The final two datasets I use in my analysis are the Survey of Industrial Research and R&D (SIRD, 1997-2007) and Business R&D and Innovation Survey (BRDIS, 2008-2015) to obtain data on firm R&D expenditures (nominal US$), R&D payroll and R&D employee counts. The SIRD and BRDIS collects firm-level annual data on R&D expenditures at total levels by type (basic research, applied research, and development), source of funding (Federal R&D funds versus company R&D funds), industry, and the number of scientists and engineers. There are 25,000-40,000 companies on average in the survey in each year and about 20 percent of the firms in the sample report positive total R&D expenditures (Foster et al., 2016). The SIRD and BRDIS are the only sources of R&D data for both private and publicly listed firms[40] and proves to be an invaluable source for a representative sample of R&D performers in the economy. However, unlike the other data sources, these innovation datasets are based on surveys and are biased towards R&D performers in terms of observable data.

LINKING DATA My analysis is based on firms that have some manufacturing presence in the US, to potentially substitute domestic production with increased input sourcing from abroad. Hence, I restrict firms with atleast one manufacturing establishment in a given year. Merging ASM and CM data along with their product trailer files to the LBD at the establishment level and aggregating it to the firm, I obtain a firm-product shipment level data for roughly 30,000 firms every year. To this, I merge the firm-product import data from LFTTD to observe the detailed set of products both produced and imported by the firms along with the nominal dollar value of imports associated with every imported product. Finally, I merge

[40] Compustat provides R&D data for only publicly listed firms.

R&D data from SIRD (1997-2007) and BRDIS (2008-2015) to construct my final unbalanced sample of approximately XX firms spanning almost two decades from 1997 to 2015.

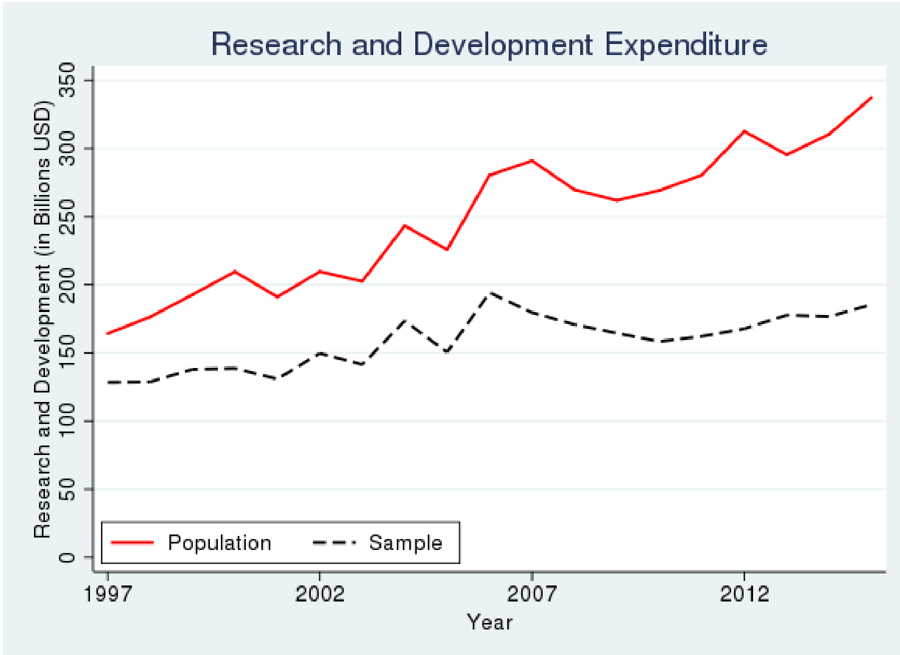


Figure 3: R&D Sample

In Figure 3, the red line denotes nominal R&D expenditures (in billions of USD) strikingly close to Foster, Grim and Zolas (2016). The black dotted line includes R&D conducted by firms in my sample.[41] Firms in my sample account for a significant share of domestic R&D in the first half of the sample period, surveyed by the SIRD. The second half, following the financial recession and introduction of BRDIS, shows that R&D performed by manufacturing firms account for a lesser fraction of national R&D.

Table 17: R&D Investment and Offshoring Participation,

	R&D Investment		
Offshoring	Yes	No	Total
Yes	46.4	12.4	58.8
No	19.4	21.7	41.10
Total	65.80	34.10	100

Table 20: Economic Effect of Offshoring on R&D Outcomes (Robustness)

Bin	Sample Share (firm-year)	Median Predicted Change
	Panel A: R&D Expenditures (1997-2015)	
% Increase in Offshoring (Annual)		
$> 50\%$	19.9%	43.72%
$25 - 49\%$	10.8%	3.77%
$5 - 24\%$	9.3%	1.45%

[41] In order to be in my sample, firms must have positive manufacturing establishments in a given year.

Table 18: Offshoring & R&D Expenditure (Full Sample)

	Log R&D Expenditure			
	(1)	(2)	(3)	(4)
Log Offshoring (3 digit)	0.070***			
	(0.020)			
Log Offshoring (4 digit)		**0.104***		
		(0.030)		
Log Offshoring (5 digit)			0.122***	
			(0.035)	
Log Offshoring (6 digit)				0.168***
				(0.049)
Estimation	IV	IV	IV	IV
Firm Fixed Effect	✓	✓	✓	✓
Industry Fixed Effect	✓	✓	✓	✓
Year Fixed Effect	✓	✓	✓	✓
Firm Controls	✓	✓	✓	✓
WES	fpc	fpc	fpc	fpc
F-state	25.81	25.70	25.22	24.87
R^2	0.881	0.881	0.880	0.879
Observations	67000	67000	67000	67000

Notes: The dependent variable is log **R&D expenditures** from 1997 to 2015. The independent variables are log "narrow" offshoring restricted at the 3-6 digit levels as that of firm shipments. *WES* denotes World Export Supply defined in subsection 3.2. Firm controls include log of capital stock, log of shipments and log of labor productivity. Industry refers to NAICS six digit classification. All regressions are weighted by SIRD/BRDIS sample weights and include a constant. All standard errors are clustered at the firm level.
******* Significant at the 1% level; ****** Significant at the 5% level; ***** Significant at the 10% level.

Table 19: Offshoring & Average R&D Earnings (Full Sample)

	Log R&D Earnings			
	(1)	(2)	(3)	(4)
Log Offshoring (3 digit)	0.187***			
	(0.014)			
Log Offshoring (4 digit)		**0.256***		
		(0.020)		
Log Offshoring (5 digit)			0.297***	
			(0.024)	
Log Offshoring (6 digit)				0.381***
				(0.033)
Estimation	IV	IV	IV	IV
Firm Fixed Effect	✓	✓	✓	✓
Industry Fixed Effect	✓	✓	✓	✓
Year Fixed Effect	✓	✓	✓	✓
Firm Controls	✓	✓	✓	✓
WES	fpc	fpc	fpc	fpc
F-state	61.34	55.74	52.22	45.64
R^2	0.794	0.787	0.781	0.765
Observations	42000	42000	42000	42000

Notes: The dependent variables are log **R&D earnings**from 1997 to 2013. The independent variables are log "narrow" offshoring restricted at the 3-6 digit levels as that of firm shipments. *WES* denotes World Export Supply defined in subsection 3.2. Firm controls include log of capital stock, log of shipments and log of labor productivity. Industry refers to NAICS six digit classification. All regressions are weighted by SIRD/BRDIS sample weights and include a constant. All standard errors are clustered at the firm level.
******* Significant at the 1% level; ****** Significant at the 5% level; ***** Significant at the 10% level.

	No of Products (HS4)		Log No of Products (HS4)	
	(1)	(2)	(3)	(4)
Offshoring Dummy	1.392***	0.506*	0.710***	0.681***
	(0.286)	(0.299)	(0.017)	(0.017)
Firm fixed effects	✓	✓	✓	✓
Year fixed effects	✓	✓	✓	✓
Size	X	✓	X	✓
R^2	0.923	0.926	0.929	0.932
Observations	67000	67000	67000	67000

	Log changes in R&D Expenditure		
	First-Stage	Reduced-Form	IV
WES Instrument	0.228***	0.017***	**0.073*****
	(0.008)	(0.006)	(0.026)
Industry × Year Fixed Effect	✓	✓	✓
Firm Controls	✓	✓	✓
WES weight	fpc	fpc	fpc
R^2	0.349	0.306	0.302
Observations	51000	51000	51000

	Log changes in R&D Earnings		
	First-Stage	Reduced-Form	IV
WES Instrument	0.286***	0.027***	**0.094*****
	(0.010)	(0.006)	(0.021)
Industry × Year Fixed Effect	✓	✓	✓
Firm Controls	✓	✓	✓
WES weight	fpc	fpc	fpc
R^2	0.321	0.420	0.415
Observations	31500	31500	31500

	Dependent variable: Log changes in R&D Expenditure (Annualized)					
	(1)	(2)	(3)	(4)	(5)	(6)
Total Offshoring (4 digit)	0.006	**0.063****				
	(0.007)	**(0.025)**				
Related Offshoring (4 digit)			-0.005	0.009		
			(0.006)	(0.016)		
Arms-Length Offshoring (4 digit)					0.006	0.067**
					(0.007)	(0.027)
Estimation	OLS	IV	OLS	IV	OLS	IV
Firm Fixed Effect	X	X	X	X	X	X
Industry × Year Fixed Effect	✓	✓	✓	✓	✓	✓
Firm Controls	✓	✓	✓	✓	✓	✓
WES		fpc		fpc-rel		fpc-arms
R^2	0.336	0.334	0.336	0.336	0.336	0.333
Observations	39000	39000	39000	39000	39000	39000