High-Skill Immigration, Offshoring R&D, and Firm Dynamics[*]

Jingting Fan[†] Eunhee Lee[‡] Valerie Smeets[§]

April 2021

Abstract

Firms' R&D activities increasingly rely on foreign inputs in the form of immigrant researchers and imported R&D services. This paper studies firms' decision on using foreign inputs in R&D and its implication on firm performance and aggregate productivity. Using administrative data from Denmark, we document two facts: that the use of foreign inputs increases firms' R&D efficiency and boosts firm performance, and that recruiting immigrants reduces the barriers that firms face in sourcing R&D services from abroad. We develop and estimate a firm dynamics model in which R&D can be done with a combination of domestic inputs, immigrant researchers, and imported R&D services. Two elements of the model—love for variety of ideas in R&D and an information channel of immigrants—rationalize the two facts and imply complementarity between different R&D inputs. Counterfactual experiments show that incorporating the use of foreign inputs in R&D is important for assessing the impacts of immigration, service offshoring, and R&D policies.

[*]We thank Lorenzo Caliendo, Kevin Lim, and Mark Roberts for helpful comments. Frank Zhang provided excellent research assistance. Any error is our own.

[†]Department of Economics, The Pennsylvania State University (e-mail: jxf524@psu.edu)

[‡]Department of Economics, University of Maryland, College Park (e-mail: eunhee11@umd.edu)

[§]Department of Economics and Business Economics, Aarhus University and CEPR (e-mail: vas@econ.au.dk)

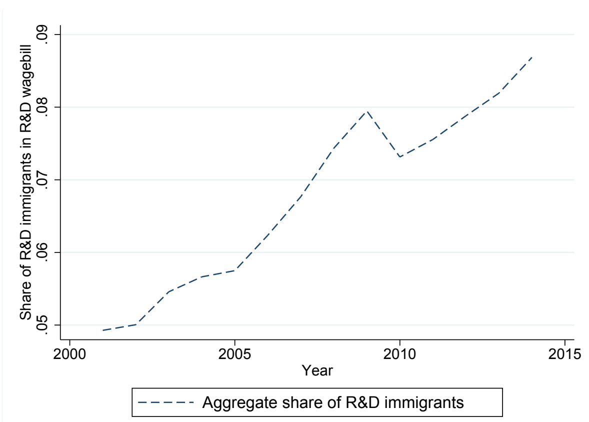
1 Introduction

Reductions in trade barriers and advances in information technology over the past few decades have made it easier for firms to source inputs from the international market. Recent studies suggest an important two-way relationship between the use of foreign intermediate inputs and productivity: on the one hand, productive firms self-select into using foreign inputs; on the other hand, access to foreign intermediate inputs improves performance both indirectly—by complementing firms' own R&D undertaking—and directly, further reinforcing the selection effect. In modeling and quantifying these channels, the literature has focused almost exclusively on the role of imported *production inputs*. Yet increasingly, firms also employ foreign *R&D inputs*, either by sourcing R&D services from abroad or by recruiting immigrant researchers to the firm.

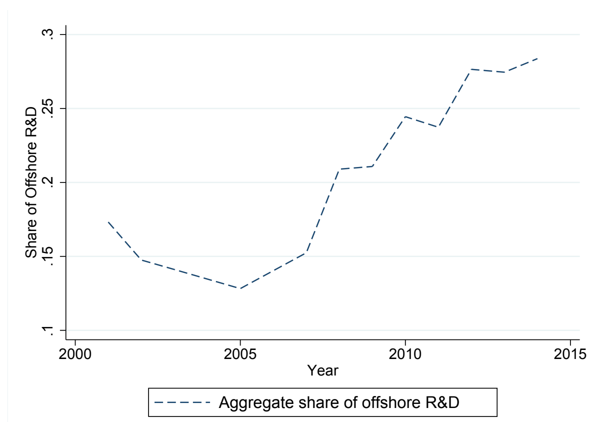
Figure 1 shows the empirical relevance of these two options for firms in Denmark, the country of focus in this paper. The left panel is immigrants' share in total research-related wage bill in Denmark. The right panel is the share of foreign sourced R&D services in total R&D expenditures. Both panels show a clear upward trend during 2001–2014, indicating increasing dependence of Danish firms on foreign inputs in R&D. By 2014, about 9% of the total R&D wage bill is spent on immigrant researchers; about 28% of the total R&D expenditures accrue to activities taking place abroad.[1]

Figure 1: Increasingly Globalized R&D Patterns of Danish Firms

(a) Immigrants among Researcher



(b) Offshore Share of R&D Expenditures



Notes: Immigrants' share in the total research wage bill is calculated using the sample of all employees in research-related occupations from the matched employer-employee data. The offshore R&D expenditures have two components: R&D services Danish firms purchase through arms' length contracts from abroad, and R&D carried out by foreign entities within the same business group of a Danish entity *for the use of that Danish entity*. The share on the right panel is the ratio between total offshore R&D expenditures and the sum of offshore R&D expenditures and R&D by Danish firms in Denmark. See Section 2 for details on data.

These two forms of foreign inputs can bring fresh ideas not seen in the domestic economy,

[1]That firms increasingly use foreign R&D inputs is hardly specific to Denmark. According to the Patent Cooperation Treaty data, between 2000 and 2010, around 10-15% of inventors in developed countries are foreign nationals (Miguelez and Fink, 2013). Relatedly, a substantial part of R&D activities in many countries are carried out by global firms outside their home countries (Fan, 2020).

boosting R&D efficiency and firm performance. In this paper we develop and quantify a dynamic heterogeneous firms model with endogenous R&D to analyze the relationship between the use of these two foreign R&D inputs and firms' productivity dynamics. The model highlights two channels. First, we posit that three inputs—domestic researchers, immigrant researchers, and R&D services from abroad—are imperfect substitutes in the R&D process, so the marginal return from R&D investment is higher, when the expenditure is spread over multiple inputs. This is the *love for variety of ideas channel*. Recruiting immigrant researchers and international sourcing of R&D services are both costly, so only a small fraction of firms take advantage of these options. However, as immigrants can provide information and know-how about foreign R&D suppliers, firms with immigrant researchers have a lower cost of sourcing R&D services from abroad, leading to a complementarity between the two options. We call this mechanism the *information channel*.

The central implications of the model are that reductions in the cost of hiring immigrant researchers indirectly encourage firms to source R&D services from abroad; conversely, facilitation of R&D offshoring also encourage firms to recruit immigrants. Moreover, because being able to combine different inputs increases the marginal return from R&D investment, both forces increase firms' participation in R&D. All these channels improve firm performance. A corollary is that to correctly assess the impact of policies on any of the three types of inputs, it is necessary to take into account the complementarity among them.

Our model is grounded in the new facts we document using rich registry data from Statistics Denmark between 2000 and 2015. We link the matched employer-employee data, which allows us to identify the occupation and immigration status of individuals, to a number of surveys and administrative data sets at the firm level, covering firms' location, accounting information, R&D status, import and export, and participation in offshoring activities. The availability of rich characteristics of both firms and workers is a unique strength of the data set and crucial for our purpose. It enables us to identify whether a firm hires immigrant researchers and how this decision correlates with the characteristics and other decisions of the firm.

We document two sets of facts from the Danish data. First, conditioning on current productivity, investment in R&D is positively correlated with the future labor productivity; more novel and important, conditioning on total R&D investment, firms engaging in more than one type of R&D inputs see higher still future productivity. This correlation is robust when we control for industry-time fixed effects and firms' participation in international markets through import and export of physical goods. The effect is present when we use standard production function estimation techniques to address simultaneity biases arising from firms' endogenous choice of production inputs. Using micro data, the literature has documented that importing intermediate production inputs increases firm productivity (see, e.g., Kasahara and Rodrigue, 2008; Goldberg, Khandelwal, Pavcnik and Topalova, 2010; Halpern, Koren and Szeidl, 2015). Our first finding suggests that the use of foreign talent or imported R&D services has an independent effect on firms' productivity dynamics above and beyond the effect stemming from the use of foreign production inputs.

Second, conditioning on industry affiliation and other characteristics, firms recruiting immigrants are more likely to conduct offshore R&D. This correlation is present at the destination-region level. For example, firms recruiting immigrants from Eastern Europe are more likely to source R&D services from Eastern Europe. We further show that this relationship is causal by using a shift-share instrumental variable (IV) for firms' recruitment of immigrants, which exploits variations across Danish regions and industries in the employment of immigrants from different source regions in 2000 and the nation-wide inflow of immigrants between 2000 and 2016. We interpret this finding as capturing that firms hiring immigrant researchers gain tacit knowledge about their home countries, which reduces the friction they face in sourcing R&D from these countries, whether it is through arm's length contracts or through in-house activities.

Motivated by these facts, we develop a model of endogenous R&D that builds on Aw, Roberts and Xu (2011), Doraszelski and Jaumandreu (2013), and Bøler, Moxnes and Ulltveit-Moe (2015). In the model, heterogeneous firms choose whether to invest in R&D, and if so, how much to invest and whether to recruit immigrant researchers or to source foreign inputs. R&D investment increases future productivity stochastically. Adopting multiple inputs at the same time is costly, but it enhances the efficiency of R&D investment and gives firms higher expected future productivity. The interpretation of this relationship is that by sourcing from a diverse source of ideas, firms can be more efficient at R&D—love for variety dating back to Ethier (1982). We also account for the information value of immigrants by allowing for reduced fixed and sunk costs of doing R&D when immigrant researchers and foreign R&D are adopted at the same time.

We uncover the importance of the love for variety of ideas in R&D and the information channel using the estimated impact of additional R&D mode on productivity: conditioning on doing R&D, firms using foreign-sourced inputs on average gain an additional 2% productivity increase. We estimate the fixed and sunk R&D cost parameters, along with the strength of the information channel, using two sources of information: the patterns of transition between different R&D status, and firms' response to an R&D subsidy program introduced in Denmark in 2011, which reduced the effective user cost of R&D by 25% for eligible firms. Our estimation results suggest that the switch from doing R&D with only native researchers to also offshoring R&D incurs a start up cost of about 10.7 million Danish Krone (DKK), or about 1.7 million USD. The presence of an immigrant researcher in the firm reduces this cost by 20%.

Through counterfactual exercises, we show that removing the information channel from our benchmark model reduces the fraction of firms conducting R&D with immigrant researchers or in offshore locations by 15 percentage points and 8.4 percentage points, respectively. This suggests that a substantial part of the value of adopting one of the foreign R&D inputs comes from the lowered cost of doing the other. The decrease in the marginal value of having an additional input type in turn leads to a decrease in firms' participation in R&D by 14 percentage points. The aggregate productivity decreases by around 0.25% as a result. We also show that the presence of the information channel amplifies the effect of policies reducing the barriers of recruiting immigrants or the cost of offshoring R&D. For example, a 50% reduction in offshoring fixed and sunk costs leads to 0.39% increase in the aggregate productivity; without the information channel, the same change in the cost of offshore R&D increases the aggregate productivity by only 0.21%.

Finally, we compare our model to a restricted model with only one R&D input to evaluate the role of the love for variety of ideas channel. When we shut down both the information channel and the love for variety of ideas channel from our benchmark model, the share of firms participating in R&D decreases even further, by 30 percentage points from the benchmark. As a result, the overall aggregate productivity decreases by 0.48% when we shut down both channels. We show that the effects of an R&D subsidy policy on firm's participation in R&D and on the aggregate productivity are significantly magnified with the love for variety of ideas channel. For example, an R&D subsidy policy which lowers fixed and sunk costs of doing any type of R&D by 50% increases the aggregate productivity by 0.77%; without the love for variety of ideas in play, the same R&D subsidy policy increases the aggregate productivity by only 0.07%.

This paper contributes to three strands of the literature. The first literature estimates the impact of imported intermediate inputs on firm performance (Amiti and Konings, 2007; Kasahara and Rodrigue, 2008; Goldberg, Khandelwal, Pavcnik and Topalova, 2010; Halpern, Koren and Szeidl, 2015; Zhang, 2017). While the literature also builds on the idea of Ethier (1982) and emphasizes a 'love-for-variety' in input use, their focus is on imported inputs for production. On the other hand, we focus on the impact of access to foreign inputs in R&D through two channels, both of which are becoming increasingly relevant as the global integration goes well beyond the exchange of goods to the exchange of ideas and movement of high-skill workers. We show that our finding is not picking up the effect of imported production inputs by directly controlling for this channel. Because R&D investment contributes to firm's knowledge capital which is carried forward to the future (subject to depreciation), improvement in R&D efficiency is accumulated and amplified over time. This also differs from the static impacts found on the bulk of the literature on imported production inputs.

Second, our firm dynamics model with endogenous R&D and estimation methodology builds on the work of Doraszelski and Jaumandreu (2013), which has also been employed by Aw, Roberts and Xu (2011) and Bøler, Moxnes and Ulltveit-Moe (2015) to study the interaction between R&D and international trade. In particular, Bøler, Moxnes and Ulltveit-Moe (2015) argues that R&D and intermediate inputs are complements and jointly impact firm performance. Relatively to existing work using this methodology, our main contribution is to open up the black box of R&D and to examine the interaction between different inputs inside the box. To incorporate and quantify different inputs into R&D, we use registry data from Denmark. Our estimated model suggests that incorporating different R&D inputs not only matters when evaluating policies on offshoring or immigration, but also matters for understanding policies focusing on R&D itself.[2]

[2]This paper is also related to a small set of recent works studying the impacts of R&D by multinational firms (Bilir and Morales, 2020; Fan, 2020) which focus primarily on how R&D in foreign affiliates of MNCs affects production in the same or nearby affiliates. The offshore R&D measure in our data includes both affiliate R&D and purchased R&D

Finally, this paper is related to a growing literature on the consequences of high-skill immigration. The literature has documented two sets of results mostly in separate settings. First, a group of existing works show that high-skill immigrants increase performance using industry, regional, and most recently, firm level variations (Markusen and Trofimenko, 2009; Peri, 2012; Ottaviano, Peri and Wright, 2018; Beerli, Ruffner, Siegenthaler and Peri, 2018). Second, existing papers have shown the presence of the information value of immigration for import, export, and offshoring. Barring a few recent exceptions (e.g. Ramanarayanan et al., 2019) that exploits firm-level status of immigrant employment, most existing work on the second channel is based on comparisons between industries or regions (Head and Ries, 1998; Rauch and Trindade, 2002; Egger, Erhardt and Lassmann, 2019; Burchardi, Chaney and Hassan, 2019; Olney and Pozzoli, 2018). Our first contribution to the literature is to document evidence of both channels in a unified setting for a specific yet important activity, R&D. Our finding suggests that one mechanism through which immigrants increase firm performance is exactly by helping the firm establish business connection at home. Our second contribution to this literature is to develop and quantify a model of firm R&D with immigrants. Compared to existing work that quantifies the impacts of immigrants using structural models (see, e.g., Bound, Braga, Golden and Khanna, 2015; Burstein, Hanson, Tian and Vogel, 2020), our model incorporates two first-order features of the data: that only the most productive firms recruit immigrants, and that immigrants and offshoring interact with each other. Both channels are important for evaluating impacts of immigration policies on firm and aggregate outcomes.

The rest of the paper is organized as follows. In Section 2, we describe the data and provide reduced-form evidence that motivates key ingredients of the model. In Sections 3 and 4, we develop and estimate the model. Section 5 reports results from counterfactual experiments. Section 5 concludes.

2 Data and Descriptive Statistics

In this section we describe our data and provide reduced-form facts on the relationship between employment of immigrant researchers, offshore R&D, and firm productivity.

2.1 Data Sources

Our analysis combines several data sets on firms and workers provided from Statistics Denmark, linked together through individual and firm identifiers. The first data set is the Integrated Data for Labor Market Research (IDA), a panel data set that covers all individuals aged between 15

from overseas vendors *for the use of the reporting entity in Denmark*, so R&D done in foreign headquarters/affiliates for local use at those foreign locations are not included in the measure. To the extent that R&D in our measure could also have an impact on the headquarters and affiliates outside Denmark, our results underestimate the impact of foreign R&D.

to 74 that are in the labor force.[3] The data we use in this paper spans from 2000 to 2015. Information on workers include nationality, age, gender, education, wage, the firm and establishment at which they work, and their occupation. Occupations are defined based on the International Standard Classification of Occupations (ISCO). As detailed in the appendix, we concord different vintages of ISCO codes and then classify all workers into R&D and non-R&D roles based on their occupation following the approach in Bernard et al. (2020). Under our classification, an occupation is R&D-related if it likely involves creative and/or technical tasks such as designing, testing, and experimenting. Example of such occupations are software developers, mechanical engineers, and technicians in chemical sciences.[4] This classification is broader than the conventional definition of R&D as activities carried out by scientists or university researchers pushing the envelop of human knowledge, but it captures that for many firms, to develop a new product, some form of experimentation and innovation is needed. Slightly abusing language, we will also refer to workers in R&D-related occupations as researchers. We define immigrant researchers as researchers whose origin country is not Denmark.

The second data set is for firms, named Regnskabsstatistik (hereafter FIRE), which is derived from the value-added tax administrative data. FIRE contains accounting information such as sales, value-added, materials, wage bill, and investments of private sector firms. Our data covers 2000 to 2015. We use fixed assets to compute capital, using the perpetual inventory method (with a discount factor of 8 % for investments).[5] We match FIRE with IDA and the matched sample accounts for 86% of the total private sector employment in Denmark.[6] Focusing on private firms with more than 10 employees, we end up with a sample of around 21,000 firms per year.

We supplement the firm-level accounting data with more detailed information on R&D from the Danish equivalent of the European Community Innovation Survey.[7] The survey aims to include the universe of R&D-active firms, so firms are by default included if they are either large (over 250 employees or revenues larger than 1 billion DKK), spend at least 5 million DKK in R&D activities, or operate in R&D-intensive industries. In addition to these firms, the survey also includes a stratified sample of smaller firms. The final sample is an unbalanced panel of around 4,000 firms per year over 2001-2015. A unique feature of this survey and crucial for our purpose, is that it contains information on not only firms' R&D expenditures within Denmark, but also their expenditures from overseas—what we call offshore R&D. The questionnaire is clear that the

[3] IDA is a snapshot of the labor market at every November, so individuals must belong to the labor force in November to be included in IDA.

[4] A drawback of this classification is that because it is based on occupation, it only captures those directly involve in related activities. If a person is heading an R&D lab as a general manager, then he/she will not be classified as R&D-related.

[5] We deflate firm-level accounting information using industry specific gross-output, value-added, materials or capital deflators. The wage bill is deflated using the consumer price index.

[6] Reporting to FIRE is only compulsory for private companies with annual turnover above 500,000 DKK and for personally own companies with annual turnover above 300,000 DKK. IDA being a snapshot of the labor market each November-year, there are also some firms in FIRE that cannot be matched with IDA. As a result, the matched sample contains the majority but not all of private-sector employment.

[7] Those surveys were initially run by a research institution (Analyseinstitut for Forskning), with Statistics Denmark taking over the responsibility of the survey in 2008.

offshore R&D expenditures reported should be for the use of the reporting entity in Denmark, so R&D done in a foreign affiliate/headquarters of a Danish firm for the affiliate/headquarters themselves—such as developing a product for production there—is not included.[8] This features differentiates the survey from other available data sets on affiliate R&D, such as the one based on the U.S. BEA, in which R&D reported in a foreign location could have been done for the use of any entities within the organization.

The R&D survey also reports information on whether a firm employees research personnel in different world regions, which will allow us to establish the relationship between the hiring of immigrant researcher and offshore R&D at the world region level. The drawback is that this firm-destination region specific measure is for in-house R&D personnel only and omits the imported R&D services through arms' length contracts. We will leverage another survey, the offshoring survey collected by Statistics Denmark in 2011 to provide additional evidence. This survey is a part of a large European collaboration through Eurostat and aims at gathering information about global value chains and international sourcing. All enterprises with 50 or more employees and a representative set of firms with 20-49 employees in manufacturing and services are sampled. A crucial piece of the information from the survey is whether and to which overseas regions a firm relocated R&D activities abroad (partly or entirely; either in house or through arms' length contracts) in 2011. We show the information channel is robust to this alternative measure.

2.2 Descriptive Statistics

Table 1 presents descriptive statistics for each of the main data sets. Panel A reports the composition of employment in Denmark by immigration status and whether the occupation is related to R&D or not. According to our classification, about 13% of workers are employed in occupations related to R&D, among which about 7% are immigrants. Not surprisingly, R&D-related workers, especially those from abroad, are more highly educated than non-R&D workers. In terms of wages, R&D related workers are paid around 250 DKK ($\approx$ 40 USD) per hour, about a third above that of non-R&D workers.

Panel B reports firms' value added per worker and their employment of R&D related workers on size. Overall, about 38% of firms have employees working on R&D related tasks. Among these firms, about a quarter hire immigrants. Both the share of firms with R&D related employees and the fraction hiring immigrants are higher for larger firms, which also tend to have higher value added per worker.

Panel C reports key statistics from the R&D survey, in which R&D status is defined based on accounting statements. Smaller firms are less likely to engage in R&D. However, conditioning on doing R&D, they tend to devote a higher fraction of revenue to it, consistent with large fixed costs in R&D. About 18% of R&D active firms also source R&D services from abroad. This share

[8] Offshore R&D include R&D expenditures either incurred by a foreign related party or outsourced through arm's length contracts. The exact wording of the questionnaire for R&D in a related party is "FoU udført af andre dele af koncernen i udlandet og anvendt internt i virksomheden", which means "R&D performed by other parts of the business group abroad and used internally in the company".

Table 1: Summary Statistics for Key Variables

Panel A: IDA	% of obs	% college+	% master +	mean hourly wage
Immigrant R&D worker	0.86	83.31	39.24	262
Immigrant non-R&D	5.86	30.33	9.76	177
Native R&D	12.08	75.67	22.49	246
Native non R&D	81.20	28.62	7.93	195
Panel B: FIRE	% of obs	mean VA/L	% R&D-related emp.	% R&D-related immi. emp
10-49	84.40	545,950	32.41	6.28
50-249	13.27	623,004	68.20	24.90
>250	2.33	682,166	88.70	59.06
All	100	559,366	38.47	9.98
Panel C: R&D Survey			among R&D Firms	
	% of obs	% R&D	mean R&D/Sales (%)	% doing off. R&D
10-49	46.38	18.84	35.29	12.26
50-249	39.18	22.79	15.63	17.03
>250	14.48	36.19	5.54	29.57
All	100	22.90	21.33	18.08
Panel D: Offshoring Survey		% of firms doing off. R&D among		mean # of dest. among off. R&D firms
	% of obs	all firm	firms with R&D immi.	
10-49	36.69	6.40	10.90	1.30
50-249	51.31	9.25	15.18	1.45
>250	12.00	20.35	30.03	2.45
All	100	9.54	17.83	1.67

Note: All statistics are based on 2011 only. The number of observation underlying Panel A is 1.8 million people; the number of firms underlying Panel B, C, and D are 20099, 3550, and 4342, respectively. Mean value added per worker and hourly wage are in DKK.

ranges from 12% among firms with fewer than 50 employees to about 30% among the biggest firms. Such heterogeneous participation rates by size suggest even for firms undertaking R&D, doing offshore R&D still incur substantial costs.

Last but not least, Panel D reports statistics from the offshoring survey. About 9.5% of all firms in the sample conduct offshore R&D. The share is higher among firms that employ immigrant researchers. Such a pattern points to potential complementarity between these two channels. Among firms that are doing offshore R&D, the average number of destination regions is 1.67; the mode of the number of destination is 1 for firms below 250 employees and 2 among the biggest group of firms.[9]

2.3 Relationship between Immigrant, Offshore R&D, and Firm Productivity

The descriptive statistics presented above point to a relationship between the use of foreign R&D inputs and firm productivity. It also indicates an interaction between the two types of foreign R&D inputs. We provide below reduced-form evidence for these mechanisms.

Sourcing of R&D inputs and labor productivity. We first present the relationship between firms' level and input choice in R&D and their performance. We estimate the following:

$$\omega_{it} = \rho\omega_{it-1} + \gamma_r \log(R\&D_{it-1}) + \gamma_{\text{off.}}\mathbb{I}(\text{off. R\&D}_{it-1}) + \gamma_{\text{immi.}}\mathbb{I}(\text{immi. R\&D}_{it-1}) + \vec{\beta}X_{it} + \phi_{j(i)t} + \zeta_{it},$$

in which ω_{it} is labor productivity (defined as value added per worker), our measure of firm performance, for firm i of industry $j(i)$ at time t.[10] ω_{it} depends on lagged performance, ω_{it-1}, firms' R&D investment, and the indicators for whether firms use foreign-source R&D inputs or immigrants in R&D. This function follows the knowledge capital model of productivity dating back to Griliches (1979), according to which ω_{it} is the knowledge capital determining firm performance. It is the sum of unappreciated knowledge capital from the previous year $\rho\omega_{it-1}$ and the new knowledge created through R&D. Our specification postulates that when firms source ideas from foreign sources, their R&D investment generates more knowledge capital: i.e., $\gamma_{\text{off.}} > 0$, $\gamma_{\text{immi.}} > 0$.

Firms' R&D investment and sourcing decisions of R&D inputs likely depend on industry characteristics and correlate with other decisions of the same firm. We capture these confounding factors through industry-time fixed effects, $\phi_{j(i)t}$, and firm decisions, X_{it}. Table 2 presents the results. Columns 1 through 4 measure $R\&D_{it-1}$ using an indicator. Column 1 suggests that doing R&D is associated with 2.9% higher productivity; conditional on doing R&D, using foreign R&D increases productivity by additional 3.7%. Column 2 shows similar is true for the use of

[9]The survey groups foreign countries into 8 groups. These groups and fraction of *offshoring firms* reporting offshoring of R&D services in each of them is: EU15 (48.7%), other European countries (29.8%), EU new member states (22%), USA and Canada (17.4%), China (16.9%), India (10.5%), other Asian countries (13.4%), and the RoW (9.1%). New member states include Poland, Hungary, Bulgaria, Romania, Slovakia, Czech Republic, Cyprus, Slovenia, Estonia, Latvia, Lithuania, and Malta. In the survey, firms can report multiple offshoring locations.

[10]Our sample is all private sector firms from FIRE matched with the R&D survey and the IDA dataset.

Table 2: Sourcing of R&D Inputs and Labor Productivity

Outcome var.	Labor productivity								
Key control	Extensive margin of R&D Status				Intensive margin: domestic R&D				Total R&D
	(1)	(2)	(3)	(4)	(5)	(5)	(6)	(7)	(8)
Lagged Labor prod.	0.677***	0.676***	0.677***	0.671***	0.676***	0.675***	0.676***	0.671***	0.671***
	(0.016)	(0.015)	(0.016)	(0.016)	(0.016)	(0.015)	(0.016)	(0.017)	(0.017)
Lagged R&D dummy	0.029***	0.027***	0.020***	0.013**					
	(0.005)	(0.005)	(0.005)	(0.005)					
Lagged log domestic R&D					0.005***	0.004***	0.004***	0.003***	
					(0.001)	(0.001)	(0.001)	(0.001)	
Lagged log total R&D									0.003***
									(0.001)
Lagged off. R&D dummy	0.037***		0.032***	0.032***	0.026**		0.024**	0.025**	0.022*
	(0.012)		(0.012)	(0.012)	(0.011)		(0.011)	(0.011)	(0.012)
Lagged immi. R&D dummy		0.036***	0.034***	0.024***		0.032***	0.031***	0.022***	0.022***
		(0.005)	(0.006)	(0.006)		(0.005)	(0.006)	(0.006)	(0.006)
Import dummy				0.043***				0.042***	0.042***
				(0.006)				(0.006)	(0.006)
Export dummy				0.018***				0.017***	0.017***
				(0.006)				(0.006)	(0.006)
Observations	33354	38175	33203	33203	33354	38175	33203	33203	33203
Industry×year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Regressions focus on for-profit firms with more than 10 employees in both FIRE and R&D surveys. Clustered standard errors (by firm) in parenthesis * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

immigrant workers in R&D. In Column 3, when both channels are added, each of them has a large, positive, and statistically significant coefficient.

A large literature in international trade has documented that importing and exporting of goods is significantly correlated with productivity. Column 4 controls for firms' importing and exporting status in period t.[11] Consistent with the literature, both variables are positive and statistically significant, but they do not diminish the role of immigrant and offshore R&D dummies. This suggests that the correlation documented for sourcing of R&D is independent of the effects through physical goods documented in the literature.

Another possibility is that firms using foreign R&D inputs are significantly out-investing other firms and part of their extra R&D investment is attributed to immigration and offshore R&D dummies. Columns 5 through 7 measure $R\&D_{it-1}$ using a continuous measure—domestic R&D expenditures. Its coefficient is statistically significant, however, the coefficient is generally not big enough for the extra R&D investment of firms using foreign inputs to drive the estimated dummies.[12] Correspondingly, the coefficients associated with offshore and immigrant R&D remain to before. Finally, in the last column, we directly control for firms' *total* R&D expenditures, the estimates are virtually the same.[13]

[11]Controlling for the lagged, instead of contemporary status gives essentially the same result.

[12]With a point estimate of 0.005, for the dummies to be entirely explained by the extra R&D investment by firms using foreign inputs, their total log R&D investment needs to be 1000 log points above that of firms not using foreign inputs.

[13]All results in Table 2 remain essentially the same, if we include additional time-varying firm controls, such as log employment.

The results in Table 2 are not causal due to a simultaneity bias in firm's input choices: as firms often observe, at least partially, their productivity, when making production and R&D decisions, our productivity measure, value added per worker, is a biased measure of firm productivity. Because $R\&D$ decision in $t-1$ depends on firm productivity in $t-1$, the bias in measuring ω_{it-1} shows up in the error term and could be correlated with firms' R&D decisions. In Section 4, we will address this concern in a theory-consistent way by imposing standard timing assumptions on firm decisions.

The nexus between immigrants and offshore R&D. Table 1 showed that firms employing immigrants in R&D related occupations are more likely to source R&D services from abroad. Here we test whether this relationship is causal. As firm-level outcomes for employment of immigrant researchers and offshoring is often confounded by other firm-level characteristics, we use variation in the origin region of immigrants by estimating the following specification:

$$\mathbb{I}(\text{off. R\&D}_{it}^{n}) = \beta\, \mathbb{I}(\text{immi}_{it}^{n}) + \tilde{\phi}_{it} + \tilde{\psi}_{tn} + \tilde{\nu}_{dn} + \tilde{\eta}_{jn} + \vec{\tilde{\beta}}\tilde{X}_{it}^{n} + \tilde{\epsilon}_{it}^{n}. \tag{1}$$

In the specification, the outcome variable is an indicator for whether firm i in industry j located in city d in Denmark employed R&D personnel located in a foreign region n in year t, constructed from the R&D survey. Between 2009 and 2015, the R&D survey has a question on which destination country groups a firm has affiliate R&D employment, which gives us an extensive margin measure of offshore R&D by firm i in destination group n. Between 2009-2012, the world is divided into 4 destination groups: EU, USA and Canada, China, and the RoW. Between 2013-2015, the world is divided into 8 groups: EU 15, EU 12, Other EU, USA and Canada, Central and South America, China, India, and RoW. We report two sets of results: based on 2009-2015 data using four destination groups; based on 2013-2015 data using 7 destination groups (we group Central and South America into the RoW as there are few variations from these countries.)

The key explanatory variable is $\mathbb{I}(\text{immi}_{it}^{n})$, an indicator for whether firm i has an employee in R&D related occupations from region n, constructed from the IDA database. $\tilde{\phi}_{it}$ and $\tilde{\psi}_{tn}$ are firm-year and foreign region-year fixed effects, which will absorb unobserved time-varying firm and destination characteristics that might drive the correlation between $\mathbb{I}(\text{off. R\&D}_{it}^{n})$ and $\mathbb{I}(\text{immi}_{it}^{n})$. $\tilde{\nu}_{dn}$ and $\tilde{\eta}_{jn}$ are city-foreign region and industry-foreign region fixed effects. Some firms might be in a Danish regions (city) that have a particularly strong connection to a foreign region, or operate in an industry where offshore region n is the most competitive offshore destination. These factors could influence both $\mathbb{I}(\text{off. R\&D}_{i}^{n})$ and $\mathbb{I}(\text{immi}_{i}^{n})$, so we control for industry-destination and city-destination fixed effects to rule out this concern. Finally, $\tilde{X}_{it}^{n}$ are control variables which capture other potential connections between firm i and region n.

Table 3 reports the results. Columns 1 is the baseline specification. With firm-year fixed effects controlled for, the specification exploits variation between different destination regions. The destination-year fixed effects control for destination-specific characteristics that affect all firms equally. The industry-destination and city-destination fixed effects address the concern that

Table 3: Immigrant Researcher and Offshore R&D

	OLS (2009-2015)		2SLS (2009-2015)			2SLS (2012-2015)		
	(1)	(2)	(3)	(4)	(5)	(7)	(8)	(9)
$\mathbb{I}(\text{immi}_{it}^{n})$	0.011***	0.011***	0.121***	0.122***	0.111***	0.114**	0.113**	0.094*
	(0.004)	(0.004)	(0.031)	(0.031)	(0.037)	(0.052)	(0.052)	(0.048)
Observations	92948	92948	92948	92948	82340	74130	74130	66234
Firm year FE	yes	yes	yes	yes	yes	yes	yes	yes
Import and export with n in t	-	yes	-	yes	yes	-	yes	yes
Destination Year FE	yes	yes	yes	yes	yes	yes	yes	yes
Industry destination FE	yes	yes	yes	yes	yes	yes	yes	yes
City destination FE	yes	yes	yes	yes	yes	yes	yes	yes
Exclude firms active in 2000	-	-	-	-	yes	-	-	yes
First stage								
$s_{j(i),d(i),2000}^{n} \cdot (L_t^n - L_{2000}^n)$			6.43***	6.40***	6.29***	9.23***	9.22***	7.21***
			(1.54)	(1.53)	(1.44)	(2.21)	(2.21)	(1.82)
Robust first-stage F	-	-	17.38	17.43	19.23	17.46	17.34	15.77

Notes: Columns 1 through 5 are specifications using 2009-2015 data, in which the world is divided into 4 regions. Columns 1 and 2 are estimated via the OLS; Columns 3 through 5 are estimated using a shifter-share instrumental variable. Columns 7 through 9 replicate Columns 3 through 5 using the 2012-2015 data with the world divided into 7 regions. Standard errors (in parenthesis) are clustered by firm in Columns 1 and 2; clustered by region-sector in Columns 3 through 9.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

certain industries or headquarter cities in Denmark have specific ties with some offshore regions. The point estimate suggests that firms hiring immigrant researcher from a foreign region n will have 1% increase in probability to conduct offshore R&D in location n.

The literature has documented strong impact of immigrant presence on regional/industry import and export. To show that our estimate is not simply picking up the relationship documented in these studies, Column 2 further controls for firm-level import and export with destination n. This barely changes the point estimate.

A remaining concern is that firms might choose to hire immigrant and conduct offshore R&D due to firm-destination specific factors unobserved to the econometrician. It is also possible that offshoring makes firms more or less likely to recruit immigrant researchers. This could bias the estimate in either directions, depending on whether these two forms of inputs are substitutes or complements in R&D. To address this concern, we employ an alternative approach, which is based on a shift-share instrumental variable that captures a firm's exposure to immigrant researchers in their location and industry. Specifically, we use the following as an IV for $\mathbb{I}(\text{immi}_i^n)$:

$$s_{j(i),d(i),2000}^{n} \cdot (L_t^n - L_{2000}^n),$$

in which $j(i)$ and $d(i)$ denote the 2-digit NACE industry and headquarter city (in Denmark) of firm i, respectively. We split Denmark into 5 geographic regions. $s_{j(i),d(i),2000}^{n}$ is the share of immigrant researchers from foreign region n in Denmark that works in city d, industry j. $(L_t^n - L_{2000}^n)$ is the increase in the total number of immigrant researchers from country n between

2000 and year t.

The relevance of this IV comes from the fact immigrants tend to be attracted to locations and industries with a high density of existing immigrants from the same origin. The exclusion restriction that we impose for identification is that, conditioning on firm fixed effects and a number of industry, city, destination country group controls, the initial distribution correlated with firm-level offshoring decisions through their employment of immigrant researchers.

Columns 3 to 5 of Table 3 display the 2SLS estimate using the data between 2009 and 2015 with the world divided into 4 regions. Column 3 is the baseline 2SLS specification; Column 4 further controls for firm-destination import and export along with firm-year fixed effects.

Our design has a reasonably strong first stage, with expected sign and K-P F statistics of 17. The point estimate is positive and statistically significant. Interestingly, the coefficient is much larger than the OLS estimate. This could be due to a reverse causality in the OLS that are corrected by the IV—firms already conducting offshore R&D might have less to gain from recruiting immigrant researchers. Alternatively, this difference could arise due to heterogeneous treatment effects. Recall the IV exploits variation in the availability of immigrants of a particular origin in the local labor market. Firms hiring in a local labor market with more immigrants of a particular origin might reap more return from the information channel, if the hired immigrant provides information about their origin country exactly due to their connection with the local immigrant community. This finding also echoes the results from the literature that uses regional and industry level data and argue for the external values of immigrants.

Because the IV is constructed using the 2000 distribution of high-skill immigrants, the exclusion restriction would be violated if the same firm hiring in 2000 also shows up in our panel and continue hiring the same set of immigrants. To rule out this concern, Column 5 includes only firms existing after 2000. The first stage F stat and the key coefficient remain essentially the same. Columns 6 through 8 replicate Columns 3 through 5 using the data from 2012-2015, a period in which we can divide the world into 7 countries. Most notably, we can separate EU, the main source of immigrant researchers in Denmark, into EU 12, EU 15, and other EU, avoiding a potential aggregation bias. The results are largely similar.

As discussed earlier, the destination group specific measure of offshore R&D from the R&D survey is for in-house employment only, but firms can also source R&D services from independent suppliers. One concern is that if in-house employment crowd out outside suppliers, then what we find in Table 3 might capture simply this substitution effect and does not imply that firms use imported R&D services more frequently. To address this concern, in the appendix, we construct a destination-specific measure of offshore R&D from the offshoring survey, which captures in-house as well as outsourced offshore R&D. Regressions based on this alternative measure is similar qualitatively to the baseline.

Taking stock, we find that foreign R&D input in the form of immigrant researchers or imported R&D services accounts for a substantial fraction of total R&D spending by Danish firms. The use of these inputs are associated with improved future firm performance above and beyond the performance increase implied by firms total R&D expenditures. Moreover, these two channels are connected to each other—the employment of an immigrant researcher causally increases the propensity of the firm to service offshore R&D.

These facts imply that policies affecting the use of one of the inputs will affect firms' use of the other inputs; both further interact with firms' R&D investment decisions. To disentangle these channels and quantify their impact on firm performance, we now turn to the model.

3 Model

In this section, we present a simple model in which firms make productivity-enhancing R&D investment by combining inputs from different sources which are imperfect substitutes. Firm's problem in our model is two-fold. The static component of the firm's problem is on the production of goods or services conditional on its current productivity and demand. The dynamic component of the firm's problem solves how to organize the R&D investment optimally by combining different types of R&D. Our model highlights the two key channels that summarize firms' incentives to use diverse R&D inputs: *love for variety of ideas*, and the *information channel*. We will describe each channel in detail in this section.

3.1 Production, Demand, and Static Profit

We start by describing the static component of the model. The production function for firm i at time t has the following form:

$$y_{i,t} = \exp\left(\omega_{i,t}\right) l_{i,t}, \tag{2}$$

where $l_{i,t}$ is the production labor at firm i, period t; $\omega_{i,t}$ denotes firm's current (log) productivity, which depends on firms' past productivity and R&D investment, to be explained below; $y_{i,t}$ is the output. Letting the price for each unit of production labor be $W_{i,t}$, the effective marginal cost for an output unit is $\frac{W_{i,t}}{\exp(\omega_{i,t})}$.[14]

Firms sell their output in a monopolistic competitive output market, characterized by the following Dixit-Stiglitz demand:

$$q_{i,t} = \left(\frac{p_{i,t}}{P_t}\right)^{\eta} Q_t, \tag{3}$$

where $q_{i,t}$ and $p_{i,t}$ are quantity and price of the variety that firm i produces; $\eta > 0$ is the demand elasticity; Q_t is the aggregate demand faced by the firm; and P_t is the corresponding ideal aggregate price index. We interpret Q_t and P_t as the effective demand encompassing the entire market faced by Danish firms, including the demand from other EU states and the rest of the world. In keeping with this interpretation, our model abstracts from firms' endogenous export decision, which could also affect their productivity (Aw et al., 2011). We motivate this simplification using

[14]When taking the model to the data, we will extend the production function to include capital and materials as inputs.

the high degree of integration of Denmark in the world economy.[15] In empirical specifications, we will control directly for exporting decisions to ensure that the main channel we document is not confounded by export. We also assume that Q_t and P_t are exogenous to firms and do not change in the counterfactual experiments we consider. This assumption is motivated by the fact that the counterfactual shocks we consider lead to less than 1 percent changes of the aggregate productivity, so those shocks alone are unlikely to drive a substantial general equilibrium response.

Firms choose l_{it} and p_{it} to maximize the static profit. Given the monopolistically competitive setting, the optimal pricing rule implies $p_{i,t} = \frac{\eta}{\eta+1}\frac{W_t}{\exp(\omega_{i,t})}$, with $\frac{\eta}{\eta+1}$ being the markup over the marginal cost. The total sales of the firm is $[\frac{\eta}{\eta+1}\frac{W_t}{\exp(\omega_{i,t})}]^{\eta+1}\frac{Q_t}{P_t^{\eta}}$. Therefore, conditional on its productivity in period t, firm i earns the following static profit at the beginning of period t:

$$\pi_t(\omega_{i,t}) = -\frac{1}{\eta}\Phi_t \cdot \exp\left((\eta+1)\ln\left(\frac{\eta}{\eta+1}\right) - (\eta+1)\omega_{i,t}\right), \tag{4}$$

in which $\Phi_t \equiv \frac{W_t^{\eta+1}Q_t}{P_t^{\eta}}$ is a shifter common to all firms, capturing the overall profitability of all firms due to wages, demand, and the market competition.

3.2 The Evolution of Productivity and Love for Variety of Ideas

Firm i's productivity evolves according to the following law of motion:

$$\omega_{i,t} = \rho\omega_{i,t-1} + \mathbf{I}_{rd_{i,t-1}>0} \cdot \gamma \cdot \log\left(rd_{i,t-1}\right) + \zeta_{i,t}, \tag{5}$$

where $\omega_{i,t-1}$ is (log) productivity of firm i in the previous period; $rd_{i,t-1}$ is firm i's total effective investment in R&D in year $t-1$ and the coefficient γ is the R&D elasticity of productivity; $\mathbf{I}_{rd_{i,t-1}>0}$ is an indicator for firms doing R&D; $\zeta_{i,t}$ is an idiosyncratic error term representing unanticipated innovation in the productivity evolution process.

Firm's total effective R&D investment is generated according to the following function:

$$rd_{i,t-1} = \left[\left(A^N\right)^{\frac{1}{\theta}}\left(rd^N_{i,t-1}\right)^{\frac{\theta-1}{\theta}} + \left(A^I\right)^{\frac{1}{\theta}}\left(rd^I_{i,t-1}\right)^{\frac{\theta-1}{\theta}} + \left(A^F\right)^{\frac{1}{\theta}}\left(rd^F_{i,t-1}\right)^{\frac{\theta-1}{\theta}}\right]^{\frac{\theta}{\theta-1}}. \tag{6}$$

In the equation, $rd^{\tilde{x}}_{i,t}$ for each $\tilde{x} \in \{N, I, F\}$ denotes input to effective R&D investment for firm i, produced by native researchers, immigrant researchers, and foreign suppliers of R&D services, respectively.[16] Using R&D services provided by foreign suppliers is effectively offshoring R&D. $A^{\tilde{x}}$ for each $\tilde{x} \in \{N, I, F\}$ denotes the efficiency of input $\tilde{x}$. $\theta > 1$ is the elasticity of substitution between the three types of R&D talent inputs. The assumption that R&D inputs from different sources are imperfect substitutes captures that each of these inputs are embedded with ideas and

[15] The openness of Denmark, measured in import plus export over GDP, is well over 100.

[16] Compared to the production function specified in equation (2), we effectively assume that the production labor and the high-skilled R&D labor are separate.

techniques from different sources, thus bringing in gains from a diversity of ideas to the R&D process. We call this channel as the *love for variety of ideas* channel through which firms have an incentive to operate diverse R&D types.

Firms can choose among different combinations of inputs in R&D. In principle, it is possible for firms to use no input from native researchers in R&D; in the data, however, this rarely happens. Thus, we assume that there are four possible combinations of R&D inputs for firms to choose from: using only inputs from native researchers (N); using inputs from both native and immigrant researchers (NI); using inputs from native researchers and foreign R&D services suppliers (NF); using all three types inputs simultaneously (NIF). We call these combinations of R&D activities as *R&D modes*. Each mode is indexed by x, where we denote the choice of doing no R&D by $x = 0$. Thus, $x \in \{0, N, NI, NF, NIF\}$.

The cost for an effective unit of R&D investment for firms choosing R&D mode $x \in \{N, NI, NF, NIF\}$, c^x, is defined as

$$
\begin{aligned}
c^N &= \left(A^N\right)^{\frac{1}{1-\theta}} p^N \\
c^{NI} &= \left[A^N\left(p^N\right)^{1-\theta} + A^I\left(p^I\right)^{1-\theta}\right]^{\frac{1}{1-\theta}} \\
c^{NF} &= \left[A^N\left(p^N\right)^{1-\theta} + A^F\left(p^F\right)^{1-\theta}\right]^{\frac{1}{1-\theta}} \\
c^{NIF} &= \left[A^N\left(p^N\right)^{1-\theta} + A^I\left(p^I\right)^{1-\theta} + A^F\left(p^F\right)^{1-\theta}\right]^{\frac{1}{1-\theta}},
\end{aligned}
$$

where $p^{\tilde{x}}$ denotes the unit price for R&D of input $\tilde{x} \in \{N, I, F\}$. Imperfect substitution between R&D inputs implies that $c^x < c^N, \forall x \in \{NI, NF, NIF\}$, i.e., given the R&D expenditure, firms using multiple sources of ideas in R&D will have larger effective R&D investment, which in turn implies that these firms will see a larger increase of productivity compared to firms using only native researchers following the law of motion in equation (5), which is consistent with patterns documented in Section 2.

3.3 Dynamic Decisions

The existence of sunk and fixed costs associated with using different R&D inputs implies that not all firms doing R&D will hire immigrants or offshore R&D. In this section, we introduce the cost parameters and describe firms' dynamic decisions. The timeline for firms' dynamic problem is as follows. At the beginning of period t, firms learn the realization of $\zeta_{i,t}$, hence their current productivity, $\omega_{i,t}$ following (5). Denote firm i's state in period t as $\mathbf{s}_{i,t} = (\omega_{i,t}, x_{i,t-1})$, in which $x_{i,t-1} \in \{0, N, NI, NF, NIF\}$ captures firms' R&D choice in period $t-1$. The R&D choice from the previous period is relevant for decision making because of R&D sunk costs which occur when firms change their R&D modes as it will be described in detail below.

Knowing $\omega_{i,t}$, firm i chooses the production level to maximizes the static profit, as described in

Section 3.1, and then determines the current period's R&D mode and total R&D expenditures. We assume that the total non-recoverable fixed and sunk costs of adopting R&D mode x is: $\tilde{F}^{x_{i,t-1},x} + \epsilon^{x}_{i,t}$. In this total cost, $\tilde{F}^{x_{i,t-1},x}$ is the systematic component that is common to all firms switching from $x_{i,t-1}$ to mode x. This cost depends on firms previous R&D status because there can be sunk cost associated with entry into a different mode, e.g., the cost of setting up a new R&D team or nestling a reliable overseas R&D supplier. $\epsilon^{x}_{i,t}$ is an idiosyncratic cost component, and we assume that $\epsilon^{x}_{i,t}$ is drawn randomly and independently (across i, t, and x) from a Type-I exteme value distribution with a scale parameter $\nu > 0$. This idiosyncratic cost can be associated with a number of factors: for example, some firms might be in a better position to recruit immigrant researchers, because it is located in a region where many immigrants live; firms that are already relying heavily on international suppliers might also find sourcing foreign R&D services relatively easier than other firms. All of these factors lead firms to make different decisions on R&D modes but are unobserved to the econometrician, so we capture these possible costs in the idiosyncratic term.

Firms observe the current draw of their idiosyncratic cost for each R&D mode and make decisions on R&D modes conditional on the draw. Let $V_t(\mathbf{s}_{i,t})$ be the expected value of a firm with current state $\mathbf{s}_{i,t}$ before the realization of $\epsilon^{x}_{i,t}$. $V(\mathbf{s}_{i,t})$ is given by:

$$
\begin{aligned}
V_t(\mathbf{s}_{i,t}) &= \pi(\omega_{i,t}) + \int \max_{x\in X} \left[V^{x}_t(\mathbf{s}_{i,t}) - \widetilde{F}^{x_{i,t-1},x} - \epsilon^{x}_{i,t} \right] d\epsilon, \qquad (7)\\
&\text{where } X \equiv \{0, N, NI, NF, NIF\}\\
&\text{and } V^{x}_t(\mathbf{s}_{i,t}) \equiv \begin{cases} \delta \cdot E_t V_t(\mathbf{s}_{i,t+1} \mid \mathbf{s}_{i,t}), \text{ for } x = 0 \\ \max\limits_{rd_{i,t}} \left\{ -rd_{i,t} \cdot c^{x} + \delta E_t V_t(\mathbf{s}_{i,t+1} \mid \mathbf{s}_{i,t}, x, rd_{i,t}) \right\}, \text{ for } x \in X \setminus \{0\} \end{cases}
\end{aligned}
$$

In Equation (7), $V^{x}_t(\mathbf{s}_{i,t})$ is the present discounted value of R&D mode x for firm i at time t; $\delta \in (0, 1)$ is the discount rate; $rd_{i,t}$ is the effective investment in R&D, aside from the fixed and sunk cost payments.

Based on the timing and distributional assumption for firms' idiosyncratic cost for R&D modes, the probability of a firm switching from an R&D mode x to an R&D mode x' is given by:

$$
m^{x,x'}_t(\mathbf{s}_{i,t}) = \frac{\exp\left(\frac{1}{\nu}\bar{V}^{x,x'}_t(\mathbf{s}_{i,t})\right)}{\sum_{x''\in X} \exp\left(\frac{1}{\nu}\bar{V}^{x,x''}_t(\mathbf{s}_{i,t})\right)},
$$

where $\bar{V}^{x,x'}_t \equiv V^{x'}_t(\mathbf{s}_{i,t}) - \widetilde{F}^{x,x'}$. Empirically, $m^{x,x'}_t(\mathbf{s}_{i,t})$ corresponds to the share among firms with productivity $\omega_{i,t}$ and mode x in period $t-1$ that switch to mode x' in period t.

We further parameterize the cost of changing R&D modes, $\widetilde{F}^{x,x'}$. Specifically, we assume that the cost $\widetilde{F}^{x,x'}$ consists of fixed operation costs of each R&D mode irrespective of firms' previous R&D status, denoted by $f^{x'}$, and a component that governs the sunk cost associated with switching between modes, denoted by $F^{x,x'}$. The total switching cost $\widetilde{F}^{x,x'}$ is the sum of the two components, i.e., $\widetilde{F}^{x,x'} = f^{x'} + F^{x,x'}$.

We stack the switching costs into a matrix,

$$\tilde{\mathbf{F}}_{5\times 5} = \mathbf{1}_{5\times 1} \cdot \mathbf{f}_{1\times 5} + \mathbf{F}_{5\times 5},$$

where the subscripts of each variable denote the dimension of the variable. Specifically, the element in the m-th row and the n-th column of matrix $\tilde{\mathbf{F}}$ is the cost of switching from mode m to mode n; $\mathbf{1}$ is a 5 by 1 vector of ones; $\mathbf{f} = (f^0, f^N, f^{NI}, f^{NF}, f^{NIF})$ is the vector of fixed operation costs; $\mathbf{F}$ is a 5 by 5 matrix of sunk cost components.

A few assumptions on $\mathbf{f}$ and $\mathbf{F}$ are in order. First, we assume that doing no R&D activity ($x = 0$) incurs neither cost, i.e., $f^0 = 0$ and $F^{x,0}$ for every x, and that there is no sunk cost if firms do not switch R&D mode, i.e., $F^{x,x'} = 0$ if $x = x'$. In addition to this assumption, we parametrize $\mathbf{F}$ by

$$\mathbf{F} = \begin{bmatrix} 0 & F^N & F^N + F^I & F^N + F^F & F^N + F^I + F^F - F^{IF} \\ 0 & 0 & F^I & F^F & F^I + F^F - F^{IF} \\ 0 & F^{I0} & 0 & F^F + F^{I0} & F^F - F^{IF} \\ 0 & F^{F0} & F^I + F^{F0} & 0 & F^I \\ 0 & F^{I0} + F^{F0} & F^{F0} & F^{I0} & 0 \end{bmatrix},$$

where each row and column corresponds to the five R&D modes $\{0, N, NI, NF, NIF\}$ in the same order, with rows indicating firms current mode x and columns indicates their next mode x'.

Components in $\mathbf{F}$ has intuitive explanations. First, F^N, F^I, and F^F terms capture the cost of setting up R&D operations carried out by native workers, immigrant workers, overseas suppliers, respectively. These are only to be paid by firms that were not using R&D inputs of each type in the previous period. Second, there is a cost associated with dropping certain types of inputs from R&D: F^{I0} for immigrant workers and F^{F0} for offshoring R&D. The idea is, when a firm stops using certain types of inputs, the rest of the R&D operation must be adjusted to accommodate the change. These adjustment cost are captured exactly through F^{I0} and F^{F0}.[17] Lastly, our *information channel* is summarized in F^{IF}. Adding offshoring R&D into firm's R&D activities may be less costly if firms do R&D with immigrant high-skilled workers at home at the same time. Immigrants at home can facilitate communications between the headquarter and offshore R&D affiliates and make information flows between the two locations easier, as our empirical evidence suggests.

The information channel, together with the love-for-variety of ideas built into the R&D function in equation (6), are the two core channels through which different R&D modes interact to determine R&D. The key parameters governing these channels are therefore, θ, the elasticity of substitution between different types of variable R&D inputs, and $\tilde{\mathbf{F}}$, the fixed and sunk cost of R&D. We will estimate these parameters in the next section.

[17]Because in the data, virtually all R&D active firms hire native researchers, we assume that when firms drop input from native researchers they shut down R&D all together. In this case, they do not need to pay the reorganization cost to keep R&D, so we assume the cost of dropping the N mode to be zero.

4 Model Estimation

In this section, we explain the procedures to estimate the model. Some parameters can be estimated without solving for firms' full dynamic optimization problem, so we estimate these parameters independently. Other parameters will be identified jointly in a nested fixed point procedure.

4.1 The Distribution of Idiosyncratic Cost Shocks

Firms' R&D choices depend crucially on the benefit of R&D, as well as the cost associated with R&D activities including their idiosyncratic cost draws for each mode. We will estimate the benefit of R&D by parameterizing firms' profit and productivity dynamics. We will use firms' patterns of transition between different R&D modes in data to identify the cost parameter **F**. The challenge, however, is that in our setting, the transition costs are in general not separately identifiable from the scale parameter ν of the distribution of the idiosyncratic cost, $\epsilon_{i,t}$. To see this, consider the share of firms switching from doing R&D with only natives to not doing R&D in period t among firms with productivity $\omega_{i,t}$, relative to the share of firms that continue on the same mode, $\frac{m_t^{N,0}(\omega_{i,t})}{m_t^{N,N}(\omega_{i,t})}$. From the expressions for $m_t^{x,x'}$ and the value functions derived above, we have the following equation:

$$\log(\frac{m^{N,0}(\omega_{i,t})}{m^{N,N}(\omega_{i,t})}) = \underbrace{\frac{1}{\nu}[c^N rd^*_{i,t} + f^N]}_{\text{R\&D expenses}} + \frac{\delta}{\nu}\underbrace{[E_t V_{t+1}(\omega_{i,t+1}|\omega_{i,t}, rd_{i,t} = 0) - E_t V_{t+1}(\omega_{i,t+1}|\omega_{i,t}, rd_{i,t} = rd^*_{i,t})]}_{\text{Improvement in continuation value from optimally chosen R\&D}}, \tag{8}$$

in which the first term captures the cost associated with continuing R&D and the second component captures the improvement in the continuation value from R&D. $rd^*_{i,t}$ is the optimally chosen quantity of the composite R&D bundle. Equation (8) shows that it is possible to rationalize firms' mode choice using $\widetilde{F}^{N,N}$ alone. More generally, when the cost matrix $\tilde{\mathbf{F}}$ is flexible enough, it will be able to explain any value of $\frac{m^{x,x'}}{m^{x,x}}$, so firms' switching of R&D modes is not informative for the identification of ν.[18]

Based on this finding, we take advantage of the introduction of an R&D subsidy regime in Denmark in 2011 as a natural experiment to identify ν, since this policy generates variation in the cost of conducting R&D. The policy specifies that for R&D active firms that are incurring a loss, 25% of their R&D cost would be rebated. With this subsidy in place, equation (8) becomes:

$$\log(\frac{m'^{N,0}(\omega_{i,t})}{m'^{N,N}(\omega_{i,t})}) = 0.75 \times \frac{1}{\nu}[rd^{*'}_{i,t} \cdot c^N + f^N] + \frac{\delta}{\nu}[E_t V'_{t+1}(\omega_{i,t+1}|\omega_{i,t}, rd_{i,t} = 0) - E_t V'_{t+1}(\omega_{i,t+1}|\omega_{i,t}, rd_{i,t} = rd^{*'}_{i,t})]. \tag{9}$$

[18]This finding is analogous to a result in the trade literature that gravity regressions alone cannot separately identify trade elasticity and trade costs.

The difference is that first, the user cost of R&D is 75% of the pre-policy level. Second, the value and policy functions could be different in general after the subsidy is introduced, so they are denoted as V' and $rd^{*'}$. We assume that firms perceive the value function in the post-policy world as similar to the one before policy, i.e., $V_{i,t}(\cdot) \approx V'_{i,t}(\cdot)$. We motivate this assumption using the uncertain and temporary nature of the policy and the qualification requirement that only loss-making firms are eligible.[19]

With this assumption, the right-hand side of the Equation (9) becomes

$$\frac{1}{\nu}\{\delta E_t V_{i,t+1}(\omega_{i,t+1}|\omega_{i,t}, rd_{i,t} = 0) + (1-0.25)f^N - \max_{rd^{*'}_{i,t}}[\delta E_t V_{i,t}(\omega_{i,t+1}|\omega_{i,t}, rd_{i,t} = rd^{*'}_{i,t}) - 0.75 \times rd^{*'}_{i,t} \cdot c^N]\}.$$

The only differences are that we now suppress the $'$ superscript in the value functions on the right hand side and highlight that $rd^{*'}_{i,t}$ is chosen to maximize the net return to R&D under the 25% subsidies. To evaluate the term involving the max operator, let τ be the R&D subsidy rate. From the Envelop theorem, we have

$$\frac{\partial \max_{rd_{i,t}}\{\delta E_t V_{i,t}(\omega_{i,t+1}|\omega_{i,t}, rd_{i,t}) - (1-\tau) rd_{i,t} \cdot c^N\}}{\partial \tau} = rd^*_{i,t} \cdot c^N,$$

where $rd^*_{i,t}$ is the optimal effective R&D when no subsidy is given. That is, the impact of R&D subsidies on the continuation value is simply equal to the direct subsidy on the intensive margin, $\tau \cdot rd^*_{i,t} \cdot c^N$, assuming the optimal R&D decision rule itself does not change.

Combining the two expressions above and using Equation (8), we have the following:

$$\begin{aligned}
&\log\left(\frac{m'^{N,0}(\omega_{i,t})}{m'^{N,N}(\omega_{i,t})}\right) \qquad (10)\\
&\approx \underbrace{\frac{1}{\nu}[rd^*_{i,t} \cdot c^N + f^N]}_{\text{R\&D expenses}} + \frac{\delta}{\nu} \underbrace{[E_t V_{i,t+1}(\omega_{i,t+1}|\omega_{i,t}, rd_{i,t} = 0) - E_t V_{i,t}(\omega_{i,t+1}|\omega_{i,t}, rd_{i,t} = rd^*_{i,t})]}_{\text{Improvement in continuation value from optimally chosen R\&D}} - \underbrace{0.25 \times \frac{1}{\nu}[rd^*_{i,t} \cdot c^N + f^N]}_{\text{R\&D subsidies}}\\
&= \log\left(\frac{m^{N,0}(\omega_{i,t})}{m^{N,N}(\omega_{i,t})}\right) - \frac{1}{\nu} \times 0.25 \times [rd^*_{i,t} \cdot c^N + f^N].
\end{aligned}$$

The last line of this equation can be interpreted as the pre-policy log odd ratio for firm with productivity $\omega_{i,t}$ net of the product of the R&D subsidy for a firm with productivity $\omega_{i,t}$ and the R&D mode switching elasticity. Equation (10) shows that with the subsidy policy in place, the change in the propensity of a loss-making firm continuing R&D is moderated by the firm's current R&D expenditures ($rd^*_{i,t} \cdot c^N + f^N$) which determines the amount of subsidies firms will

[19] From the perspective of firms qualifying for this subsidy in year t, they would qualify in the next year only if both of the following conditions hold: i) the subsidy policy is still active; ii) they continue to be in a loss position; and iii) they are actively doing R&D. Given the uncertainty in policy and the potential upside risk of R&D-active firms, firms likely do not anticipate all three conditions to hold in the future. We also note that assuming $V_{i,t}(\cdot) \approx V'_{i,t}(\cdot)$ does not mean that firms always perceive their values to be the same as before. The correct interpretation of the assumption is that, conditioning on productivity and the mode of R&D chosen by the firm, the perceived value is the same as before.

receive, and ν, which governs firms' responsiveness to the level of subsidy. In the model, firm characteristics are uniquely determined by their productivity and participation in R&D. Equation (10) suggests that we can determine ν by checking if loss making firms are less likely to quit doing R&D after the policy is introduced and if the decrease in their propensity to quit is higher for firms investing more in R&D.

Table 4 reports the transition of loss-making firms in 2011 (before the R&D subsidy policy, left panel) and 2012 (after the policy, right panel). There are around 240 firms in both periods. Before the subsidy was enacted, 26% of firms doing R&D in the previous period stopped doing so; in 2012, when the subsidies took in effect, only 14% quit, which is consistent with the goal of the subsidy policy trying to reduce the number of firms giving up on R&D activities.

Table 4: R&D Decisions of Loss-Making Firms

Loss-making firms in 2011				Loss-making firms in 2012			
		2011				2012	
		N	0			N	0
2010	N	53	19	2011	N	72	12
	0	16	155		0	19	135

To control for differences in firm's productivity, we split observations from 2011 and 2012 into $k = 1, ..., K$ bins based on firm's labor productivity. We then estimate the following linear probability function

$$\mathbf{I}_{i \text{ quits R\&D in } t} = \beta_0 + \beta_k \mathbb{I}_{\omega_{i,t} \in k} + \ln(\text{emp}_{i,t-1}) + \epsilon_{i,t},$$

where $\mathbf{I}_{i \text{ quits R\&D in } t}$ is a dummy variable that takes a value of 1 if firm i quits doing R&D in year t; and β_k is a fixed effect for all observations belonging to the k-th productivity bin. To rule out the possibility that the estimate is driven by changing macroeconomic conditions which may affect big and small firms differentially, we always control for firm size, measured as lagged log employment.

Columns 1 through 3 in Table 5 report the estimation results from this specification, with the number of productivity bins ranging from 5 to 20. Column 4 is the result when we control for a flexible function of firm productivity instead of the bin dummy. All columns give similar estimates: after the policy is introduced, firms are about 12% less likely to quit doing R&D activities. Columns 5 through 8 provide a placebo test by focusing on the profit-making firms, which were not eligible for the subsidy. It shows that among this group of firms, there was no change in their propensity to quit doing R&D. So the results in Columns 1 to 4 are not due to a change in the macroeconomic conditions between 2011 and 2012 that affected all firms simultaneously.

The estimates reported in Table 5, however, do not directly translate into the structural pa-

Table 5: R&D Subsidy and Firm R&D: Linear Regression

	Loss-making firms				Placebo: profitable firms			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Policy dummy	-0.125***	-0.114***	-0.126***	-0.116***	-0.023	-0.023	-0.022	-0.019
	(0.040)	(0.040)	(0.042)	(0.039)	(0.024)	(0.024)	(0.024)	(0.024)
Observations	414	414	414	414	1100	1100	1100	1100
Bin FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of bins	5	10	20	-	5	10	20	-
Controls	size	size	size	size+prod.	size	size	size	size+prod.

Notes: Robust standard errors in parenthesis * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

rameter of interest, $\frac{1}{\nu}$. Columns 1 through 4 of Table 6 estimate a logistic specification with the independent variable being log of R&D expenditure, which corresponds to an elasticity specification of Equation (10).[20] The estimated coefficients are all around -0.14. Since our subsequent productivity estimation will focus on the manufacturing sector, Columns 5 through 8 also report results for only manufacturing firms. The sample shrinks by two thirds, but the estimated coefficients are relatively similar to the ones from all industries. Our preferred specification in Column 7 of the Table gives -0.232 as the estimate of elasticity. To translate this number into a value for the semi-elasticity given in Equation (10), $-\frac{0.25}{\nu}$, we assume that the firm with the median R&D expenditure in the sample has an elasticity of -0.232. With the median R&D expenditure being 2 million DKK in the corresponding sample, we have $-\frac{0.25}{\nu} = \frac{-0.232}{2}$, which implies $\nu = 2.16$.[21]

Table 6: R&D Subsidy and Firm R&D: Logistic Regression

	All industries				Manufacturing Only			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Policy*Log R&D	-0.147***	-0.139***	-0.163***	-0.136***	-0.199***	-0.209***	-0.232***	-0.195***
	(0.0371)	(0.0383)	(0.0422)	(0.0365)	(0.0740)	(0.0787)	(0.0882)	(0.0729)
Observations	363	363	352	363	107	107	103	107
Bin FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of bins	5	10	20	-	5	10	15	-
Controls	size	size	size	size+prod.	size	size	size	size+prod.

Notes: Sample focus on loss-making firms only. R&D expenditures are measured in million DKK, or around 160 thousan USD. Robust standard errors in parenthesis * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

[20] Alternatively, we could use the level of R&D expenditure as the explanatory variable to directly estimate $-\frac{0.25}{\nu}$. We do not adopt this specification because the distribution of R&D expenditure is highly skewed. Specifications with the level of expenditure as the explanatory variable are thus heavily influenced by a small number of big firms. Therefore, we choose to estimate a logit specification, before converting it to level. Nevertheless, we note that using a level specification gives qualitatively similar results.

[21] Note that the scale of ν depends on the unit of measurement. $\nu = 2.16$ corresponds to the specification with the currency unit in million DKK.

4.2 R&D and Productivity Evolution

The second set of parameters, including ρ, γ, σ_ζ (standard deviation of ζ), $A^{\tilde{x}}$, and $p^{\tilde{x}}$, governs the evolution of firm productivity. Letting $e_{i,t-1}$ denote the R&D expenditure that firm i spends within Denmark, the period-t productivity of the firm i doing any R&D in the previous period is given by:

$$\begin{aligned}
\omega_{it} =& \rho\omega_{it-1} + \gamma\log(e_{it-1}) - \gamma\log(c^N) \\
&+ \begin{cases} \zeta_{i,t}, & \text{if } x_{i,t-1} = N \\ \gamma(\log(c^N) - \log(c^{NI})) + \zeta_{i,t}, & \text{if } x_{i,t-1} = NI \\ \gamma\theta[\log(c^N) - \log(c^{NF})] + \zeta_{i,t}, & \text{if } x_{i,t-1} = NF \\ \gamma[\theta(\log(c^{NI}) - \log(c^{NIF})) + (\log(c^N) - \log(c^{NI}))] + \zeta_{i,t}. & \text{if } x_{i,t-1} = NIF \end{cases}
\end{aligned} \tag{11}$$

Each line of this equation indicates one mode of doing R&D. For example, for firms using mode N, the effective R&D bundle is $\log(rd_{i,t-1}) = \log(e_{it-1}) - \log(c^N)$. Suppose we regress the current productivity on lagged productivity, total domestic R&D spending within Denmark, and a dummy variable for the use of immigrant researchers. Then, the recovered coefficient on the dummy variable would have a structural interpretation of $\gamma[\log(c^N) - \log(c^{NI})] > 0$, which is straightforward from comparing the first and the second cases of Equation (11). This coefficient implies that the same domestic R&D expenditures will contribute more to productivity growth, if they are spent on R&D activities done by both domestic and immigrant researchers.

The third line of the equation for the case of $x_{i,t-1} = NF$ specifies the productivity of firms using imported R&D services, i.e., offshoring R&D. Conditional on the total R&D spending *within* Denmark, these firms will see an additional benefit of $\gamma\theta\big(\log(c^N) - \log(c^{NF})\big)$ on their productivity. Note the coefficient before the term is $\gamma\theta$ instead of γ as in the case of NI mode. This difference is due to the fact that we first took out the R&D investment within Denmark before specifying the four cases in the equation, so the adjustment $\gamma\theta\big(\log(c^N) - \log(c^{NF})\big)$ captures the combined effect of the following two forces. First, conditional on R&D investment in Denmark, firms that are also offshoring will make additional spending on R&D; second, given the total spending on R&D, the effective bundle of R&D will be larger. Finally, the last line of the equation is for firms doing all three R&D modes. Because $c^{NI} > c^{NIF}$ and $c^N > c^{NI}$, the systematic component for this case is also positive and greater than the systematic component for firms in the NI mode.

Motivated by Equation (11), we specify the following regression:

$$\omega_{it} = \overline{\omega} + \tilde{\rho}\omega_{it-1} + \tilde{\gamma}_0\log(e_{it-1}) + \tilde{\gamma}_1\mathbb{I}(x_{i,t-1} = NI) + \tilde{\gamma}_2\mathbb{I}(x_{i,t-1} = NF) + \tilde{\gamma}_3\mathbb{I}(x_{i,t-1} = NIF) + \zeta_{i,t}, \tag{12}$$

in which $\overline{\omega}$ is a constant productivity shifter that captures the cost of domestic R&D which is common to all firms. $\tilde{\rho}$ and $\tilde{\gamma}_m$ for each $m = 0, 1, 2, 3$ are structural parameters corresponding to the parameters in Equation (11).[22] Equation (12) also clarifies that $\tilde{\gamma}_m$ for each $m = 1, 2, 3$, summarizes the benefit of having access to additional sources of ideas, due to either a large elasticity of substitution θ or a large decrease in effective cost (i.e., $\log c^N - \log c^{NI}$). These parameters contain all information about $A^{\tilde{x}}$ and $p^{\tilde{x}}$ that firms need to know when deciding whether to adopt an additional input in R&D. As an econometrician, once we have recovered these reduced-form estimates, we can plug them back in Equation (11) and treat it as the law of motion of productivity firms expect when solving their R&D decisions.

There are two challenges in estimating Equation (12). First, $\omega_{i,t}$ is unobserved to the econometrician. Second, firm's R&D are often measured with errors. More specifically, given the finite sample and the fact that only few firms are in the NF mode, we might not be able to estimate all reduced-form parameters in Equation (12) precisely, which poses a serious challenge as we need to feed the estimates into the model. To overcome the first challenge, we will follow a two-step approach to recover $\omega_{i,t}$ from firms' input and output and estimate the parameters of the production function. To overcome the second challenge, we will estimate an auxiliary regression that are more robust to measurement errors and discipline the model by matching the estimates from the auxiliary regression. We explain these procedures in great detail below.

Recall that in the previous section, we have assumed that firms use only labor. In recovering ω_{it} and the process characterizing productivity evolution, we introduce capital, labor, and materials as input in production to be consistent with the data. We then estimate the elasticities of these inputs using a control function approach, which has been extended to estimate the return to R&D (see e.g., Doraszelski and Jaumandreu, 2013; Aw et al., 2011; Bøler et al., 2015).

Formally, the log of the revenue of a firm is:

$$\tilde{y}_{i,t} = \beta_k \cdot \tilde{k}_{i,t} + \beta_l \cdot \tilde{l}_{i,t} + \beta_m \cdot \tilde{m}_{i,t} + \omega_{i,t} + e_{i,t}, \tag{13}$$

where $\tilde{y}_{i,t}$, $\tilde{k}_{i,t}$, $\tilde{l}_{i,t}$, and $\tilde{m}_{i,t}$ denote the log of revenue, capital, labor, and materials, respectively, for firm i in period t. In practice, $\omega_{i,t}$ captures both firms' productivity and the quality of their product.[23] As standard in the literature, we assume that materials are a static input, chosen after firms observe $\omega_{i,t}$ in period t. $e_{i,t}$ is the measurement error in revenue or productivity shocks that are realized after firms' decisions are already made. Because materials are chosen with the knowledge of $\omega_{i,t}$, an OLS estimation of Equation (13) is subject to a simultaneity bias. Following the insight of Levinsohn and Petrin (2003) and Ackerberg et al. (2015), we adopt a two-step control function approach as detailed below.

Step 1. The first step is to come up with a control of $\omega_{i,t}$ using observables. Noting that $\tilde{m}_{i,t}$ is chosen given $\tilde{k}_{i,t}$, $\tilde{l}_{i,t}$, and $\omega_{i,t}$, we write the material use as a general function $\tilde{m}_{i,t} = m_t(\omega_{i,t}, \tilde{k}_{i,t}, \tilde{l}_{i,t}, z_{i,t})$, where $z_{i,t}$ are other characteristics of firm i at time t that might affect its

[22] Concretely, $\tilde{\rho} = \rho$; $\tilde{\gamma}_0 = \gamma$; $\tilde{\gamma}_1 = \gamma(\log(c^N) - \log(c^{NI}))$; $\tilde{\gamma}_2 = \gamma\theta[\log(c^N) - \log(c^{NF})]$; $\tilde{\gamma}_3 = \gamma[\theta(\log(c^{NI}) - \log(c^{NIF})) + (\log(c^N) - \log(c^{NI}))]$.

[23] When both quality and productivity follow the same persistence and respond to R&D in the same way, it is without loss of generality to encapsulates them in one term.

material use. Following Doraszelski and Jaumandreu (2013), we include $p^l_{i,t}$, the labor input price of firm i, in $z_{i,t}$, which allows us to account for heterogeneous wages between firms(Fox and Smeets, 2011). Finally, a robust finding from the trade literature is that access to imported intermediates has a positive impact on firm performance (e.g., Amiti and Konings, 2007; Halpern et al., 2015). In our context, access to imported materials can have an impact on material use by reducing the effective price of intermediate goods. To account for this channel, in a robustness we will include firms' importing status in $z_{i,t}$.

Assuming monotonicity of $m_t(\omega_{i,t}, \tilde{k}_{i,t}, \tilde{l}_{i,t}, z_{i,t})$ in $\omega_{i,t}$ for given values of $\tilde{k}_{i,t}$, $\tilde{l}_{i,t}$, and $z_{i,t}$, we can invert $m(\cdot, \tilde{k}_{i,t}, \tilde{l}_{i,t}, z_{i,t})$ to express $\omega_{i,t}$ as an (unknown) time-varying function of capital, labor, material use, and $z_{i,t}$: $\omega_{i,t} = \tilde{\omega}_t(\tilde{k}_{i,t}, \tilde{l}_{i,t}, \tilde{m}_{i,t}, z_{i,t})$. By substituting this into Equation (13), we get

$$\begin{aligned} y_{i,t} &= \beta_k \tilde{k}_{i,t} + \beta_l \tilde{l}_{i,t} + \beta_m \tilde{m}_{i,t} + \tilde{\omega}_t(\tilde{k}_{i,t}, \tilde{l}_{i,t}, \tilde{m}_{i,t}, z_{i,t}) + \epsilon_{i,t} \\ &\equiv h_t(\tilde{k}_{i,t}, \tilde{l}_{i,t}, \tilde{m}_{i,t}, z_{i,t}) + \epsilon_{i,t}. \end{aligned}$$

We specify $h_t(\cdot)$ as a cubic function of capital, investment, employment, wage, and a yearly dummy and estimate $\hat{h}_{i,t}$ using OLS in the first step.

Step 2. With $\hat{h}_t(\cdot)$ from the first step, we can express $\omega_{i,t} = \hat{h}_t(\cdot) - \beta_k \tilde{k}_{i,t} - \beta_l \tilde{l}_{i,t} - \beta_m \tilde{m}_{i,t}$. By substituting this into Equation (12), we obtain:

$$\begin{aligned} \hat{h}_{i,t} - \beta_k \tilde{k}_{i,t} - \beta_l \tilde{l}_{i,t} - \beta_m \tilde{m}_{i,t} &= \overline{\omega} + \rho \cdot (\hat{h}_{i,t-1} - \beta_k \tilde{k}_{i,t-1} - \beta_l \tilde{l}_{i,t-1} - \beta_m \tilde{m}_{i,t-1}) \\ &+ \tilde{\gamma}_0 \log(e_{i,t-1}) + \tilde{\gamma}_1 \mathbb{I}(x_{i,t-1} = NI) + \tilde{\gamma}_2 \mathbb{I}(x_{i,t-1} = NF) + \tilde{\gamma}_3 \mathbb{I}(x_{i,t-1} = NIF) + \zeta_{i,t}. \end{aligned} \tag{14}$$

As discussed in Ackerberg et al. (2015), in a setting like ours, material usage does not have independent variation, once capital, labor, and productivity are controlled. We adopt the first-order approach (see, e.g., Gandhi et al., 2020) to estimate β_m using the average revenue shares of materials. We then plug in $\hat{\beta}_m$ to Equation (14) and estimate the remaining parameters using the Generalized Method of Moments (GMM).

Note that $k_{i,t}$ and $e_{i,t-1}$ are determined before the innovation of firm productivity, $\zeta_{i,t}$, realizes, so they are independent of the error term. Labor use, on the other hand, may react to $\zeta_{i,t}$. Since $\tilde{l}_{i,t-1}$ and $\tilde{k}_{i,t-1}$ are chosen before $\zeta_{i,t}$ is known, they can serve as instrumental variables (IV) for $\tilde{l}_{i,t}$. We further allow for $\overline{\omega}$ to vary by industry. To this end, we add industry dummies to the specification and use these dummies as their own IVs.

Table 7 reports the results. Columns 1 through 3 control for the intensive margin of R&D expenditures. The first column is our baseline specification. The estimated coefficient for the intensive margin of R&D is statistically significant, but is fairly small, as in the OLS. This could be due to measurement errors in this margin, or masking heterogeneous impacts by industry and firm size. Both findings are in line with the estimate based on the Norwegian data by Bøler et al. (2015).[24] More important to our purpose, we find statistically significant and economically

[24] Bøler et al. (2015) find a larger R&D coefficient for big firms and a negative coefficient for small firms, so the average return is of around a similar magnitude. If we split firms in our sample by size and estimate for heterogeneous

sizable impacts of adopting NI and NIF roles. The coefficient for the NF dummy is slightly negative and rather imprecisely estimated. This is likely due to the fact that in the sample, only a small number of firms adopt the NF mode.

Table 7: R&D and Productivity Evolution

	Control for Log Domestic R&D			Control for R&D Indicator					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
ρ	0.468***	0.479***	0.479***	0.468***	0.479***	0.478***	0.464***	0.473***	0.472***
	(0.151)	(0.128)	(0.128)	(0.145)	(0.120)	(0.116)	(0.150)	(0.121)	(0.102)
$\log(e_{i,t-1})$	0.002***	0.002***	0.002***						
	(0.001)	(0.001)	(0.001)						
$\mathbb{I}_{i,t-1}(i \in N)$				0.010***	0.009***	0.009***	0.011***	0.010***	0.010***
				(0.004)	(0.003)	(0.004)	(0.004)	(0.003)	(0.003)
$\mathbb{I}_{i,t-1}(i \in NI)$	0.021*	0.021**	0.021**	0.023*	0.022**	0.022**			
	(0.008)	(0.007)	(0.007)	(0.008)	(0.007)	(0.007)			
$\mathbb{I}_{i,t-1}(i \in NF)$	-0.005	-0.006	-0.006	-0.002	-0.003	-0.003			
	(0.008)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)			
$\mathbb{I}_{i,t-1}(i \in NIF)$	0.041**	0.042**	0.042**	0.046**	0.048**	0.048**			
	(0.014)	(0.013)	(0.013)	(0.015)	(0.013)	(0.013)			
$\mathbb{I}_{i,t-1}(i \in NI \bigcup NF \bigcup NIF)$							0.022*	0.022**	0.022**
							(0.007)	(0.006)	(0.006)
Input elasticities									
β^l	0.452	0.450	0.450	0.453	0.451	0.451	0.456	0.454	0.454
	(0.010)	(0.010)	(0.009)	(0.010)	(0.010)	(0.009)	(0.011)	(0.010)	(0.009)
β^k	0.110	0.109	0.109	0.111	0.110	0.110	0.112	0.111	0.111
	(0.016)	(0.013)	(0.012)	(0.016)	(0.013)	(0.012)	(0.015)	(0.013)	(0.010)
β^m	0.454	0.454	0.454	0.454	0.454	0.454	0.454	0.454	0.454
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Industry fixed effects	yes	yes	yes	yes	yes	yes	yes	yes	yes
Import dummy		yes	yes		yes	yes		yes	yes
Lag export dummy			yes			yes			yes
Number of observations	9320	9320	9320	9320	9320	9320	9331	9331	9331

Notes: Manufacturing firms with 10+ employees only. β^m is computed as the revenues share of materials. Bootstrapped standard errors in parenthesis using 200 replications. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The second column includes the import dummy in both the control function $h(\cdot)$ and in the firm's law of motion as in Equation (14). This specification allows importing to play two roles, both of which have been documented in the literature: directly improving firm productivity; changing firms' pattern of material use for any given productivity. Column 2 of the table suggests that our main finding is not due to a potential correlation between using different R&D inputs and importing. Column 3 allows export to affect productivity, which is essentially a 'learning by exporting' effect well-documented in the literature, and we find that most coefficients do not change. The lower panel of the table reports the estimates for the input elasticities. All estimates are reasonable and stable across specifications.

Auxiliary Regression. As discussed earlier, due to measurement errors and the small number of observations with the NF mode, we might not be able to estimate the coefficients precisely. We thus consider an auxiliary regression, in which we replace the intensive margin measure of R&D with a dummy and group some modes of R&D together. For transparency, we introduce

returns, we obtain similar findings.

our auxiliary regression in two steps.

Specifically, in Columns 4 through 6 of Table 7, we include an indicator for conducting R&D, instead of $\log(e_{i,t-1})$, in the regression. The coefficients for other variables do not change substantively, and the coefficient for R&D dummy is around 1%. This R&D dummy captures the average impact of $\log(e_{i,t-1})$ for the firm with the average log R&D expenditure.

In Columns 7 through 9 of Table 7, we further replace the mode-specific dummies for NI, NF, NIF with a joint dummy, which takes the value of one if any of these modes are on. Loosely speaking, this dummy captures the frequency-weighted average value of $\tilde{\gamma}_1$, $\tilde{\gamma}_2$, and $\tilde{\gamma}_3$. Our preferred estimate, in Column 9 of the Table 7, suggests that on average, doing R&D increases productivity by 1%; adopting additional modes other than native workers adds another 2%.

We take three moments from Column 9, the estimated autocorrelation, the R&D dummy, and the diverse mode dummy. We supplement these moments with three additional moments: 1) the average R&D expenditure share on immigrants among NI firms (s^I_{NI}); 2) the average R&D expenditure share on foreign suppliers among NIF firms (s^F_{NIF}); 3) the standard deviation of log firm sales ($sd(\tilde{y})$). We stack these moments in a vector, denoted by $\hat{\alpha}_{6\times 1}$, which we will use in an indirect inference procedure to pin down $\tilde{\rho}$, $\tilde{\gamma}_i, i = 0, 1, 2, 3$, and σ_ζ.

Formally, for any given values of $\tilde{\rho}$, $\tilde{\gamma}_i, i = 0, 1, 2, 3$, and σ_ζ, we simulate the model and use the simulated data to generate targeted regression coefficients $\hat{\alpha}$ corresponding to Column 9 of Table 7 and calculate the standard deviation of log sales and moments on firms' R&D shares (s^F_{NIF} and s^I_{NI}). We then search over the space of $\tilde{\rho}$, $\tilde{\gamma}_i, i = 0, 1, 2, 3$ and σ_ζ so that these model-generated moments are exactly the same as the target. Because the model-implied coefficients on $\hat{\alpha}$ depend on the share of firms doing R&D and their distribution over different modes, which depends on the fixed and sunk costs of the model, we combine the indirect inference with the estimation of the other parameters of the model, as will be described in the next subsection.

4.3 Joint Estimation of the Remaining Parameters

Aside from the parameters governing the law of motion of firm productivities, the remaining parameters are: the aggregate demand shifter Φ_t, and the fixed and sunk costs of various R&D modes, $\widetilde{\mathbf{F}}$. We estimate these parameters jointly. Specifically, for given parameters, we solve firms' optimization problem and simulate the steady state of the model. We then compare the moments of the steady state distribution to corresponding moments in the data. We search for the parameters so the distance between the model moments and their empirical counterparts are minimized. Below we describe the moments that identify each parameter and our estimation procedure.

Aggregate Demand Shifter. Φ_t is the aggregate demand shifter for all firms. Since we focus on the steady state of the model, we assume that Φ_t is a constant and picks Φ_t so that the median sales of firms in our model match that in the data. For the firms in the transition matrix, the target moment is 144 million DKK (or 23 million USD).

Fixed and Sunk Costs of R&D Modes. $\widetilde{\mathbf{F}}$ directly determines the probability that a given firm switches from one mode of R&D to another. We can thus pin down $\widetilde{\mathbf{F}}$ using the observed transition matrix. Formally, let $m^{x,x'}$ be the fraction of mode x firms in a period that chooses mode x' in the next period. This leaves us with 20 moments in the transition matrix to pin down 10 parameters in the transition matrix because 5 of them are redundant as probabilities should sum up to 1 for each origin mode. We weight the transition by the number of firms in each mode. In calculating the transition matrix, we focus on the same sample as in the GMM estimation to ensure consistency.

Estimation Procedures. We classify the parameters to be estimated into two categories. The first category—parameters for productivity and firm size—are just identified, with the same number of moments as the parameters. The second category of the parameters, $\widetilde{\mathbf{F}}$, are over identified. To maintain a tight connection between the parameters and the moments that identify them, our estimation solves the following constrained optimization problem:

$$\min_{\theta\in\Theta} \sum_{x,x'} n(x) \cdot \left(m^{x,x'}(\theta) - \hat{m}^{x,x'}\right)^2$$
$$s.t.\ \alpha(\theta) = \hat{\alpha},$$

in which variables with a hat are empirical moments; variables without a hat are model-implied values under the parameter $\theta \in \Theta$. For example, $\hat{\alpha}$ is the three estimated coefficients reported in Column 9 of Table 7—the estimated autocorrelation, the R&D dummy, and the diverse mode dummy—which are our targeted moments, and $\alpha(\theta)$ are the same estimated coefficients with the simulated data with parameters θ. The constraints in the problem guarantee that all just-identified moments are matched; the objective function then minimizes the deviation between the model and the empirical transition properties, in which $\hat{m}^{x,x'}$ is weighted by the share of firms in mode x in the data, denoted by $n(x)$. The constraint ensures that the coefficients from the auxiliary regressions and three additional moments on firm size and mode-specific R&D shares meet the data.

Estimation Results and Fit of the Model. Panel B of Table 8 summarizes the moments identifying each parameters and the values of these parameters. Among other things, $\rho = 0.47$, similar to the empirical counterpart. Consistent with an attenuation bias for intensive margin R&D expenditures, $\tilde{\gamma}_0$ is larger than the estimated coefficient in Column 3. Perhaps for this reason, the estimated $\tilde{\gamma}_1$ and $\tilde{\gamma}_3$ are both smaller than in the empirical specification.

Our estimate of $\tilde{\gamma}_i, i = 0, 1, 2, 3$ implies that $\theta = 1.36$. This elasticity highlights that different sources of R&D inputs are not very substitute. This finding is a direct result of the large gains from using multiple modes in the auxiliary regression.

Table 9 Panel A reports the empirical transition matrix and the model counterpart. The model is able to fit the patterns in the data well. The mean difference between the model and the data are in the order of 0.02. The fit of the row '*NF*' is worse than other, perhaps because there are very few firms in '*NF*' mode, so these moments are weighted less. The last row of Panel A reports the mode distribution in the data as well as implied by the steady state of the model.

Table 8: Summary of Structural Parameters

Parameters	Descriptions	Source/Target	Value
A. Estimated Independently/calibrated			
ν	idiosyncratic cost in R&D	Table 6	10
η	demand elasticity	Aw et al. (2011)	-6.56
B. Jointly Estimated			
ρ	prod. autocorr	Table 7 Column 9	0.469
$\tilde{\gamma}_i, i = 0,1,2,3$	return to R&D	Table 7 Column 9 + shares	$\tilde{\gamma}_0 = 0.0034$, $\tilde{\gamma}_1 = 0.0009$, $\tilde{\gamma}_2 = 0.0074$, $\tilde{\gamma}_3 = 0.0066$
σ_ξ	sd. of the inno. term in prod.	sd(log($sales$))	0.1897
Φ	agg. demand	median sales: 144 million DKK	-
$\widetilde{F}^{x,x'}$	Fixed and sunk costs in R&D	Table 9	Table 9

Given that the model matches the transition matrix reasonably well, it is perhaps unsurprising that the model matches this piece of the data well, even though they are not directly targeted.

Panel B of Table 9 reports the estimated coefficients. The left side of the panel is the total cost of transition. Two remarks on the composite cost. First, the diagonal elements are generally smaller than the other parameters in the same row, suggesting that sunk cost play an important role. In terms of quantitative magnitude, we find that the start up cost of doing R&D domestically only is around 8 million DKK. This accounts for about 80% of the average R&D expenditures among firms underlying the transition matrix sample (manufacturing firms with 10+ employees).

Second, note that $\tilde{F}^{N,NF} > \tilde{F}^{NI,NIF}$, i.e., for a firm having immigrant researchers, the cost of switching on offshore R&D is around 20% lower. This reflects the information channel documented in our reduced-form section. The right side of the panel is the break down by mode-specific fixed cost and the sunk cost associated with mode switching. We can see that $\tilde{F}^{N,NF} - \tilde{F}^{NI,NIF} = F_{IF} + f_F - f_{IF}$. Two thirds of the information channel comes from the fixed cost component and the remaining component comes from the sunk cost component. Our estimation also suggests that there are considerable cost in shedding the immigrant mode; the cost of shedding offshore is essentially zero.

5 Counterfactuals

In this section, we simulate our benchmark calibrated model and compare the results with simulations based on various counterfactual scenarios. First, we study the role of the two main channels of our model, the information channel and the love for variety of talent channel, in shaping the distribution of R&D modes, aggregate productivities, and the share of R&D expenditures from different modes. We can quantify the role of each channel by shutting down each channel one by one from our benchmark model and comparing the simulation results to the results from our benchmark model. We then turn to the effect of counterfactual policy changes which resemble real-world policy changes, such as immigration policies for high-skill talents,

Table 9: Transition Matrix and Cost Estimates

Panel A: Transition probability and stead state distribution: model versus data

	0		N		NI		NF		NIF	
	Data	Model	Data	Model	Data	Model	Data	Model	Data	Model
0	0.889	0.895	0.063	0.076	0.033	0.017	0.005	0.007	0.011	0.006
N	0.277	0.257	0.591	0.557	0.082	0.118	0.042	0.039	0.008	0.028
NI	0.117	0.109	0.059	0.056	0.679	0.700	0.012	0.003	0.133	0.132
NF	0.135	0.119	0.385	0.258	0.034	0.055	0.405	0.394	0.041	0.174
NIF	0.049	0.039	0.011	0.020	0.251	0.252	0.022	0.026	0.667	0.662
SS dist.	0.583	0.572	0.136	0.136	0.168	0.173	0.021	0.021	0.092	0.099

Panel B: Recovered cost matrix and its breakdowns

$\widetilde{F}^{x,x'}$	0	N	NI	NF	NIF	fixed	f^N	f^{NI}	f^{NF}	f^{NIF}
0	0	7.261	12.895	17.777	21.650	cost	0.187	0.000	3.949	2.956
N	0	0.187	5.821	10.703	14.576					
NI	0	3.220	0.000	13.736	8.755	sunk	F^N	F^I	F^F	F^{IF}
NF	0	0.187	5.821	3.949	8.776	cost	7.074	5.821	6.754	0.954
NIF	0	3.220	0.000	6.982	2.956		F^{I0}	F^{F0}		
							3.033	0.000		

offshore R&D liberalization, and R&D subsidies. Lastly, we quantify the role of the information channel as well as the love for variety channel as the transmission mechanisms of the effect of each policy change on various aggregate outcomes.

5.1 The Role of the Information Channel and the Love for Variety of Talents Channel

In our model, firms have an incentive to hire immigrant talents for two main reasons: the information channel and the love for variety of talents channel. We study the role of each channel on the distribution of R&D modes and the aggregate productivity in this subsection, where the aggregate productivity is calculated by the average of productivity of firms, weighted by their sales.

The information channel, through which immigrants make offshoring less costly by facilitating communications between the headquarter and the offshore R&D affiliates, operates through both the fixed operation cost and the sunk switching cost of our model. In the previous section, we find that the calibrated value of f^{NIF} is lower than that of f^{NF}, which suggests that having immigrant researchers at home makes running offshore R&D affiliates less costly. In addition, F^{IF} in the matrix of sunk switching costs specifically captures the information channel as the positive value of this parameter lowers the cost of switching to the R&D mode involving both immigrants and offshoring R&D compared to the cost of adding offshoring R&D without immigrants. Therefore, in order to shut down the information channel from our model, we consider an alternative model where we set f^{NIF} equal to the calibrated value of f^{NF} as well as $F^{IF} = 0$.

We simulate the benchmark model with the estimated and calibrated parameters from the previous section as well as the alternative model with the aforementioned parameter restrictions

Table 10: Distribution of R&D Modes and Aggregate Productivity: Benchmark vs. No Information Channel

	(a) Share of firms by mode (%)		(b) Share of total R&D expenditure (%)		(c) Aggregate (log) productivity	
R&D modes	Benchmark	No info.	Benchmark	No info.	Benchmark	No info.
No R&D	56.44	70.23	–	–	0.231	0.236
N	13.79	15.09	29.98	48.96	0.251	0.260
NI	17.55	10.83	38.17	35.49	0.244	0.260
NF	2.13	2.12	5.62	8.58	0.342	0.365
NIF	10.09	1.73	26.23	6.97	0.326	0.358
All	-	-	-	-	0.254	0.252

Notes: Columns marked as "No info." report the results from the alternative model specification where there is no information channel operating. Panels (a) and (b) report the share (%) of firms by R&D mode and the share (%) of R&D expenditure from each mode, respectively. Panel (c) reports the average of log productivity for all firms that belong to each R&D mode, weighted by each firm's sales. The last row for Panel (c) reports the aggregate log productivity for all firms regardless of the R&D decision, weighted by each firm's sales.

to shut down the information channel. Table 10 compares the distribution of firms across R&D modes and the share of R&D expenditures by mode at the steady state between the benchmark model and the alternative model without the information channel. As shown in panel (a) of the table, without the information channel operating, the share of firms not doing any R&D activity is 14 percentage points larger. Even among the firms engaging in R&D activities, significantly less firms do R&D with all three types of talents. The drop is most pronounced for the R&D mode NIF because shutting down the information channel eliminates the synergy that immigrant researchers bring in by making offshoring R&D easier. This difference in composition across R&D modes translates into the different share of R&D expenditures by mode as reported in panel (b) of the table. The share of R&D expenditure by firms doing the most diversified R&D drops from 26% to 7% when we shut down the information channel.

Difference in R&D patterns has an important implication for the aggregate productivity as well as the productivity distribution across different R&D modes. In Table 10, we also report the aggregate log productivity for all firms and for firms with each R&D mode from each model specification in panel (c). We compute the aggregate log productivity by taking an average of firm's log productivity, weighted by individual firm's simulated sales. The aggregate productivity for all firms is 0.25% higher when our information channel is at play. This result can be explained by the compositional shift between R&D modes documented in Table 10. Without the information channel at play, significantly more firms decide not to do any R&D activity and much less firms do R&D activities with multiple talent types because it is much more costly to do so compared to the benchmark model. As a result, only a small number of the most productive firms elect to do diversified R&D. This compositional shift increases the aggregate productivity for all R&D modes in the specification without the information channel, while it decreases the overall aggregate productivity.

We then explore the role of the love for variety of talents channel by shutting it down from the benchmark model. To shut down this channel, we assume that the fixed operation costs for

Table 11: Comparison between the Benchmark and the Model without the Love for Variety of Talents Channel

	(a) Share of firms by mode (%)		**(b) Aggregate (log) productivity**	
R&D modes	Benchmark	No love for variety	Benchmark	No love for variety
No R&D	56.44	87.31	0.231	0.244
R&D	43.56	12.69	0.278	0.280
All	-	-	0.254	0.249

Notes: Columns marked as "No love for variety" report the results from the alternative model specification where there is no love for variety of talents channel operating. Panel (a) reports the share (%) of firms by R&D decision. Panel (b) reports the average of log productivity for all firms that belong to each group, weighted by each firm's sales. The last row for Panel (b) reports the aggregate log productivity for all firms regardless of the R&D decision, weighted by each firm's sales.

diversified R&D modes (NI, NF, and NIF) are all infinite. Therefore, firm's effective choice set includes only an option to do no R&D at all and an option to do R&D activities with only one type of talent. By shutting down the love for variety of talents channel from the benchmark model, we are thus effectively shutting down the information channel as well. In other words, this exercise shuts down both the information channel and the love for variety of talents channel.

Table 11 compares the share of firms doing R&D and the aggregate productivity between the benchmark model and the model without the love for variety of talents channel. Since this alternative specification effectively shuts down both the information channel and the love for variety of talents channel, we see much larger differences from the benchmark model compared to the changes reported from the previous exercise where we shut down only the information channel. The share of firms doing R&D is 30.78 percentage points smaller when neither channel is operating. Panel (b) shows that this compositional shift affects the aggregate productivity, as the overall aggregate productivity is smaller with neither channel operating. Like the result from the alternative specification where only the information channel is shut down, the aggregate productivity for each group of firms goes up without the love for variety of talents channel, because the compositional shift between no R&D and R&D is the key adjustment margin.

In summary, we find that both the information channel and the love for variety of talents channel play key roles in firms' decisions on R&D activities. Both channels increase firms' incentives to do R&D activities in general, and having both channels operating leads to more diversified R&D activities and higher aggregate productivity.

5.2 High-Skill Immigration and Offshore R&D Policies

High-skill immigration and offshore R&D are both affected a great deal by changes in individual country's policies. Those policy changes often trigger heated debates on their economic and social impacts. We use our model framework to study the effects of such policy. First, we analyze the effect of a change in the immigration policy for high-skilled workers, where it becomes less costly for firms to newly hire high-skilled immigrants for their R&D activities. Specifically, we lower F^I in the matrix of sunk switching costs by 50% from the baseline estimate. Second, we study the effect of a policy which brings firms increased access to high-skilled talents overseas. For example, when a country with a large pool of high-skilled workers opens its door to attract foreign firms to set up R&D affiliates, this policy will be reflected as a lower sunk switching cost F^F in our model. For this experiment, we lower F^F by 50%. In addition to analyzing the effects of these policies on firm's decision of R&D modes and aggregate productivity, we also explore the role of the information channel as a transmission mechanism for the effect of such policy changes. For this study, we simulate each counterfactual scenario with the alternative model without the information channel as discussed in the previous subsection, and compare the magnitude of changes in outcome variables of interest between the benchmark and the alternative models.[25]

Table 12 compares the share of firms and the share of R&D expenditures by R&D mode between the benchmark model and each of counterfactual cases with policy changes. Both the immigration and the offshoring R&D policies lower the sunk switching cost of the corresponding mode by 50% respectively as described above. The immigration policy that lowers the cost of newly hiring immigrant high-skilled workers significantly increases the share of firms hiring immigrants for their R&D activities by 10.69 percentage points. In addition, the share of firms both hiring immigrants and offshoring R&D increases by 3.27 percentage points as well because firms hiring immigrants will find offshoring R&D easier due to the information channel. The offshoring R&D policy which lowers the cost of newly establishing R&D offshoring affiliates overseas increases the share of firms offshoring R&D–in particular, the share of firms doing R&D by hiring all three types of talents (NIF) increases by 13.88 percentage points. This result is also due to the information channel. Firms hiring immigrants at home face a significantly lower cost of offshoring R&D as we show with the calibrated model. Therefore, those firms can take advantage of the lower switching cost of offshoring R&D much more than the firms that do not have immigrants at home.

Another interesting finding is that not only the share of firms doing all three types of R&D but also the share of firms hiring immigrants but not offshoring R&D also increases in response to the lower cost of offshoring R&D. This result can be also explained by the information channel. Since we assume that the policy change is permanent, all firms realize that establishing foreign R&D affiliates has permanently become easier due to this policy change. Firms that hire immigrants will benefit more from this policy in a sense that it is easier for them to take advantage of the lower cost of offshoring R&D than the firms that do not have immigrant talents. In addition, not all firms are productive enough to switch directly to the mode NIF. Therefore, some firms have an incentive to switch to NI in anticipation of the possibility of enjoying the information channel in the future.

Both policies change the distribution of R&D expenditures across firms with different R&D

[25] We do not compare the effects of these policy changes between the benchmark model and the model without the love for variety of talents channel, because the distinction between different type of talents matters only when the love for variety of talents channel is active. We discuss the quantitative role of the love for variety of talents channel further with a different counterfactual policy change in the next subsection.

Table 12: Changes in R&D and Productivity– Benchmark versus Counterfactual Policy Changes

	(a) Share of firms by mode (%)			(b) Share of total R&D expenditure (%)			(c) Aggregate (log) productivity		
R&D modes	Benchmark	Immigration	Offshoring	Benchmark	Immigration	Offshoring	Benchmark	Immigration	Offshoring
No R&D	56.44	44.07	30.57	-	–	–	0.231	0.226	0.221
N	13.79	14.99	11.70	29.98	25.19	15.63	0.251	0.242	0.235
NI	17.55	24.97	30.09	38.17	42.44	40.58	0.244	0.243	0.231
NF	2.13	2.61	3.67	5.62	5.27	5.81	0.342	0.326	0.315
NIF	10.09	13.36	23.97	26.23	27.11	37.97	0.326	0.324	0.309
All	-	-	-	-	-	-	0.254	0.255	0.258

Notes: Columns marked as "Immigration" report the results from the counterfactual scenario about the immigration policy, and the columns markets as "Offshoring" report the results from the counterfactual change of the offshoring policy, as described in the text. Panels (a) and (b) report the share (%) of firms by R&D mode and the share (%) of R&D expenditure from each mode, respectively. Panel (c) reports the average of log productivity for all firms that belong to each R&D mode, weighted by each firm's sales. The last row for Panel (c) reports the aggregate log productivity for all firms regardless of the R&D decision, weighted by each firm's sales.

mode. The immigration policy increases the share of R&D expenditure by firms hiring immigrants in the total R&D expenditure of the economy by 5.15 percentage points, and the offshoring policy increases the share of R&D expenditure by firms offshoring R&D by 11.93 percentage points. In the case of the offshoring policy, the increase is predominantly driven by firms doing all types of R&D activities because they are the most productive firms to begin with and are able to enjoy the benefit of the information channel from the immigrant talent they have.

We compare the aggregate productivities for all firms and by each R&D mode for each counterfactual scenario. As shown in Table 12, both policy changes increase the aggregate productivity for the overall economy: by 0.1% from the immigration policy and by 0.4% from the offshoring policy. Changes in the aggregate productivity by R&D mode from each policy change show the similar pattern to the case of the no information channel discussed in the previous subsection in the sense that compositional shifts toward more diversified R&D are the main reason why the aggregate productivity for the overall economy increases. The aggregate productivity by each R&D mode all decreases from both immigration and offshoring policies, which confirms the increase of the aggregate productivity driven by compositional shifts. This result is because we change the sunk switching cost in the counterfactual scenario for both types of policy experiments.

Lastly, we explore the role of our information channel in transmitting the effects of these policy changes based on a difference-in-differences analysis. Specifically, we compare the change of the share of firms by R&D mode and the change in the aggregate productivity from each counterfactual policy change between the benchmark model and the model without the information channel. We shut down the information channel as we discussed in the previous subsection. Table 13 reports changes in the share of firms per R&D mode and the aggregate productivity between the benchmark and each counterfactual scenario and compares them with the changes under the alternative model without the information channel. For both policies, the information channel clearly magnifies adjustments firms between R&D modes. In other words, the effect of both policies giving a better incentive for diversified R&D activities is much larger when the information channel is in play. As shown in Panel II of Table 13, this magnification effect of

Table 13: Counterfactual Changes with or without the Information Channel

	Immigration policy					Offshoring R&D policy				
	No R&D	N	NI	NF	NIF	No R&D	N	NI	NF	NIF
I. Changes in the share of firms by mode (pp)										
with the information channel	-12.37	1.20	7.42	0.48	3.27	-25.88	-2.09	12.54	1.54	13.88
without the information channel	-12.55	1.53	9.47	0.21	1.34	-12.51	1.66	4.51	3.14	3.20
II. Changes in the aggregate productivity (overall, %)										
with the information channel	0.14					0.39				
without the information channel	0.11					0.21				

Notes: The results reported in Panel II. are for the entire economy.

the information channel leads to a larger increase of the aggregate productivity for the overall economy in response to the positive shock on the immigration and offshoring R&D policies.

5.3 R&D Subsidies in the Age of Globalized R&D

Many countries adopt various policies to better incentivize firms to invest more in R&D, and Denmark is no exception. These policies are typically in the form of direct subsidies, tax credit, etc. For example, Denmark has offered firms in a tax loss position a refund for the deficit related to their R&D expenses since 2012. More specifically, eligible companies receive tax credits which correspond to 22% of the deficit related to their R&D expenditures up to DKK 5.5 million per year. Since this policy does not make a distinction between the type of R&D expenditures that firms claim, we design our policy experiment in such a way that all types of R&D fixed operation and sunk switching costs go down equally by 50% from the baseline level.

We then explore the quantitative role of the love for variety of talents channel in shaping the effect of such R&D subsidy policy on the distribution of firms across different types of R&D and the aggregate productivity. We shut down the love for variety of talents channel from our benchmark model by assuming that the fixed and sunk costs of any diversified R&D mode is infinite. Since the information channel matters only when there are diversified R&D modes, the alternative model without the love for variety of talents channel effectively shuts down the information channel as well.

Table 14 summarizes the changes in the share of firms and the share of R&D expenditures by R&D mode for the benchmark case and the counterfactual scenario of R&D subsidy. The R&D subsidy policy which lowers all kinds of R&D fixed and sunk costs by 50% significantly increase the share of firms doing R&D activities overall. Within the set of firms that elect to engage in R&D activities, this R&D policy induced more firms doing diversified R&D. Among the four types of R&D modes, the share of NIF firms increased the most, by 28.86 percentage points, due to the R&D subsidy policy. We also see a similar pattern for the share of total R&D expenditure by each R&D mode. As it was the case for the immigration policy and the offshoring

Table 14: Changes in R&D and Productivity – Benchmark versus R&D Subsidy Policies

	(a) Share of firms by mode (%)		**(b) Share of total R&D expenditure (%)**		**(c) Aggregate (log) productivity**	
R&D modes	Benchmark	R&D subsidy	Benchmark	R&D subsidy	Benchmark	R&D subsidy
No R&D	56.44	15.18	-	–	0.231	0.199
N	13.79	12.35	29.98	13.02	0.251	0.214
NI	17.55	19.92	38.17	21.47	0.244	0.221
NF	2.13	13.60	5.62	16.67	0.342	0.284
NIF	10.09	38.95	26.23	48.84	0.326	0.293
All	-	-	-	-	0.254	0.262

Notes: Panels (a) and (b) report the share (%) of firms by R&D mode and the share (%) of R&D expenditure from each mode, respectively. Panel (c) reports the average of log productivity for all firms that belong to each R&D mode, weighted by each firm's sales. The last row for Panel (c) reports the aggregate log productivity for all firms regardless of the R&D decision, weighted by each firm's sales.

policy in the previous subsection, this compositional shift toward more diversified R&D leads to a much higher overall aggregate productivity for the economy, while the aggregate productivity for each R&D mode goes down. An R&D subsidy policy lowering all fixed and sunk costs of R&D activities by 50% increases the overall aggregate productivity by 0.8%.

We can study the role of the the love for variety of talents channel in the effect of such R&D policy on the distribution of firms across R&D modes and the aggregate productivity by shutting down the channel from our benchmark model. For this alternative specification, we assume that the fixed and sunk costs of doing diversified R&D are all infinite. In other words, we allow for only a single type of R&D. As in the case of the immigration policy and the offshoring policy in the previous subsection, we do a difference-in-differences analysis re-evaluating the effect of the R&D subsidy policy under this alternative model specification shutting down the love for variety of talents channel.

Table 15 compares the effect of the same R&D subsidy policy on the share of firms doing R&D activities and the aggregate productivity between the benchmark model and the alternative model without the love for variety of talents channel. With this channel operating, much less firms choose to do R&D activities. The share of firms not doing R&D decreases by 27.6 percentage points less in response to the counterfactual R&D subsidy policy when we shut down the love for variety of talents channel from our benchmark model. Since much less firms choose to do R&D, the increase of the overall aggregate productivity from this R&D policy is cut by half without the love for variety of talents channel in play. In summary, allowing for a variety of R&D types for firms magnifies the positive effect of R&D subsidy policies on the aggregate economy.

6 Conclusion

In the integrated world economy, firms source their input from suppliers around the world. While the impact of access to intermediate inputs has been well documented, firm's use of foreign inputs in R&D is not well understood. Using unique data from Denmark that links workers to firms and covers firms' domestic use of foreign R&D, we study empirically and quantitatively

Table 15: Counterfactual Changes from R&D Subsidy with or without the Love for Variety of Talents Channel

	No R&D	N	NI	NF	NIF
I. Changes in the share of firms by mode (pp)					
Benchmark	-41.26	-1.42	2.37	11.47	28.86
Without the information or the love for variety channel	-13.67	13.67*			
II. Changes in the aggregate productivity (overall, %)					
Benchmark	0.14				
Without the information or the love for variety channel	0.07				

Notes: The number with * in Panel I denotes the change in the share of firms doing any R&D in percentage points, because there is no distinction between R&D modes when the love for variety of talents channel is shut down. The results reported in Panel II. are for the entire economy.

firms' decision to hire immigrant researchers and to use import R&D services, and the joint impacts of these decisions on firm performance. We find that hiring immigrant researchers reduces the barriers firm face in sourcing R&D services from abroad. Moreover, firms' R&D investment generates a higher return, when they are able to use either of the two foreign inputs, which we interpret as a 'love-for-idea' effect in R&D. We develop and estimate a model of firm dynamics with endogenous R&D choices. Counterfactual experiments using the model suggest an important role of the information channel through immigration for firm productivity; because of the complementarity among the three types of R&D inputs, policies affecting one R&D input will also have an amplified effect on firm- and aggregate outcomes.

This paper is a step towards a better understanding of firms' global organization of R&D. One of the strength of our data is that it includes firms' use of R&D services produced abroad, which allows us to interpret foreign sourced R&D as input into the R&D of firms in Denmark. Many large global firms carry out R&D in foreign locations to be used for local production, which our measure do not capture. To be consistent with our data, we control for these forces using firms' import and export in the empirical analysis, but abstract from modeling affiliate R&D in the quantitative section. Modeling and measuring the distribution of R&D expenditures among affiliates and their uses is an important venue for future research.

References

Ackerberg, Daniel A, Kevin Caves, and Garth Frazer, "Identification properties of recent production function estimators," *Econometrica*, 2015, *83* (6), 2411–2451.

Amiti, Mary and Jozef Konings, "Trade liberalization, intermediate inputs, and productivity: Evidence from Indonesia," *American Economic Review*, 2007, *97* (5), 1611–1638.

Aw, Bee Yan, Mark J Roberts, and Daniel Yi Xu, "R&D investment, exporting, and productivity dynamics," *American Economic Review*, 2011, *101* (4), 1312–44.

Beerli, Andreas, Jan Ruffner, Michael Siegenthaler, and Giovanni Peri, "The Abolition of Immigration Restrictions and the Performance of Firms and Workers: Evidence from Switzerland," *American Economic Review (Forthcoming)*, 2018.

Bernard, Andrew B, Teresa C Fort, Valerie Smeets, and Frederic Warzynski, "Heterogenous Globalization: Offshoring and Reorganization," Technical Report, NBER Working Paper 26864 2020.

Bilir, L Kamran and Eduardo Morales, "Innovation in the global firm," *Journal of Political Economy*, 2020, *128* (4), 1566–1625.

Bøler, Esther Ann, Andreas Moxnes, and Karen Helene Ulltveit-Moe, "R&D, international sourcing, and the joint impact on firm performance," *American Economic Review*, 2015, *105* (12), 3704–39.

Bound, John, Breno Braga, Joseph M Golden, and Gaurav Khanna, "Recruitment of foreigners in the market for computer scientists in the United States," *Journal of labor economics*, 2015, *33* (S1), S187–S223.

Burchardi, Konrad B, Thomas Chaney, and Tarek A Hassan, "Migrants, ancestors, and foreign investments," *The Review of Economic Studies*, 2019, *86* (4), 1448–1486.

Burstein, Ariel, Gordon Hanson, Lin Tian, and Jonathan Vogel, "Tradability and the Labor-Market Impact of Immigration: Theory and Evidence From the United States," *Econometrica*, 2020, *88* (3), 1071–1112.

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

Egger, Peter H, Katharina Erhardt, and Andrea Lassmann, "Immigration and firms' integration in international production networks," *European Economic Review*, 2019, *111*, 1–34.

Ethier, Wilfred J, "National and International Returns to Scale in the Modern Theory of International Trade," *The American Economic Review*, 1982, *72* (3), 389–405.

Fan, Jingting, "Talent, Geography, and Offshore R&D," *Unpublished Manuscript*, 2020.

Fox, Jeremy T. and Valerie Smeets, "Does Input Quality Drive Measured Differences in Firm Productivity?*," *International Economic Review*, 2011, *52* (4), 961–989.

Gandhi, Amit, Salvador Navarro, and David A Rivers, "On the identification of gross output production functions," *Journal of Political Economy*, 2020, *128* (8), 2973–3016.

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

Griliches, Zvi, "Issues in assessing the contribution of research and development to productivity growth," *The bell journal of economics*, 1979, pp. 92–116.

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

Head, Keith and John Ries, "Immigration and Trade Creation: Econometric Evidence from Canada," *Canadian journal of economics*, 1998, pp. 47–62.

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

Levinsohn, James and Amil Petrin, "Estimating production functions using inputs to control for unobservables," *The Review of Economic Studies*, 2003, *70* (2), 317–341.

Markusen, James R and Natalia Trofimenko, "Teaching locals new tricks: Foreign experts as a channel of knowledge transfers," *Journal of Development Economics*, 2009, *88* (1), 120–131.

Miguelez, Ernest and Carsten Fink, "Measuring the International Mobility of Inventors: A New Database," *WIPO Economics & Statistics Series No.8*, 2013.

Olney, William W and Dario Pozzoli, "The Impact of Immigration on Firm-level Offshoring," *Review of Economics and Statistics*, 2018, pp. 1–19.

Ottaviano, Gianmarco IP, Giovanni Peri, and Greg C Wright, "Immigration, trade and productivity in services: evidence from UK firms," *Journal of International Economics*, 2018, *112*, 88–108.

Peri, Giovanni, "The Effect of Immigration on Productivity: Evidence from US States," *Review of Economics and Statistics*, 2012, *94* (1), 348–358.

Ramanarayanan, Ananth et al., "Immigrants and Exports: Firm-level Evidence from Canada," *Unpublished Manuscript*, 2019.

Rauch, James E and Vitor Trindade, "Ethnic Chinese networks in international trade," *Review of Economics and Statistics*, 2002, *84* (1), 116–130.

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

A Data and Empirics

A.1 Data Sources

IDA. The data reports the nationality of every individuals which allows us to identify immigrants, defined as first generation immigrants or foreign individuals who are not born in Denmark. This data set covers around 2.3 million individuals and 130,000 firms per year. Many firms in Denmark are small, restricting our sample to firms with at least 10 employees leaves us with around 2 millions workers and 27,000 firms per year.

As workers can have multiple employment spells over the year, we only keep the primary job held by workers in November. Individual occupation classifications are based on the Danish version of the International Standard Classification of Occupations (ISCO). As the occupation classification changes in 2008, we concord codes over time following Mons (2018). We identify occupations related to R&D activities using a subset of the major groups 2 and 3 of the ISCO code, following Bernard et al. (2020).

A.2 Additional Evidence

The information channel based on the offshoring survey. The offshoring survey is available for 2011, so we use a cross-sectional specification. To address potential endogeneity concerns, we follow the shift-share instrument for identification, as in the baseline specification. Table 16 reports the results. We find that the recruitment of immigrant researchers has a strong positive impact on the offshoring of R&D to their origin country groups. Column 2 shows that this result is independent of firm's importing and exporting connection to the country group. Column 3 exclude firms active in 2000 to ensure that the shares in the shift-share design is not driven by the decision of the firms in our regression sample. All three columns give similar estimates. The point estimates are larger than in Table 3. This is likely due to that the definition of offshore R&D here includes both in-house and arms' length activities.

B Additional Derivations

Driving Equation (4)

$$\begin{aligned}
\pi(\omega_{i,t}) &= -\frac{1}{\eta}[\frac{\eta}{\eta+1}\frac{W_t}{\exp(\omega_{i,t})}]^{\eta+1}\frac{Q_t}{P_t^{\eta}} \\
&= -\frac{1}{\eta}\frac{W_t^{\eta+1}Q_t}{P_t^{\eta}} \cdot \exp\left((\eta+1)\ln\left(\frac{\eta}{\eta+1}\right) - (\eta+1)\omega_{i,t}\right), \\
&\equiv -\frac{1}{\eta}\Phi_t \cdot \exp\left((\eta+1)\ln\left(\frac{\eta}{\eta+1}\right) - (\eta+1)\omega_{i,t}\right),
\end{aligned}$$

Table 16: The Information Channel: Cross-Sectional Evidence from the Offshoring Survey

Dependent variable: $\mathbb{I}(\text{offshore R\&D}_i^n)$			
	(1)	(2)	(3)
$\mathbb{I}(\text{immi}_i^n)$	0.359***	0.365***	0.445***
	(0.133)	(0.134)	(0.168)
Observations	32760	32760	28672
Firm FE	yes	yes	yes
Industry destination FE	yes	yes	yes
City destination FE	yes	yes	yes
Import and export with n	-	yes	yes
Exclude firms active in 2000	-	-	yes
First stage			
$s^n_{j(i),d(i),2000} \cdot (L^n_{2011} - L^n_{2000})$	1.891***	1.887***	1.636***
	(0.486)	(0.477)	(0.445)
Robust first-stage F	14.786	14.797	8.788

Notes: Results are based on the 2011 offshoring survey. Instrument are constructed based on the growth in overall immigrants from different regions to Denmark, interacted with their initial (2000) distribution in different Danish industries and regions. Columns 1 and 2 include firms in the sample; Column 3 excludes firms present in 2000. Standard error clustered by region-sector.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$