

**Managers as Knowledge  
Carriers – Explaining Firms’  
Internationalization Success  
with Manager Mobility**

*Philipp Meinen, Pierpaolo Parrotta, Davide Sala, Erdal Yalcin*

## **Impressum:**

CESifo Working Papers

ISSN 2364-1428 (electronic version)

Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH

The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute

Poschingerstr. 5, 81679 Munich, Germany

Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email [office@cesifo.de](mailto:office@cesifo.de)

Editors: Clemens Fuest, Oliver Falck, Jasmin Gröschl

[www.cesifo-group.org/wp](http://www.cesifo-group.org/wp)

An electronic version of the paper may be downloaded

- from the SSRN website: [www.SSRN.com](http://www.SSRN.com)
- from the RePEc website: [www.RePEc.org](http://www.RePEc.org)
- from the CESifo website: [www.CESifo-group.org/wp](http://www.CESifo-group.org/wp)

# Managers as Knowledge Carriers – Explaining Firms’ Internationalization Success with Manager Mobility

## Abstract

How does “what managers know” affect firm performance on international markets? This question is of considerable importance in the international economic literature. Answering it will be key for comprehending the way firms’ varying performance on international markets is shaped by the human factor. This paper proposes managerial mobility as an integral part of such an answer. Catering products to an international customer base entails a learning process, which, to a large degree, stems from the experience of doing it. Therefore, different employers immensely contend for managers’ highly valuable export experience. As managers can accept better and better positions from several offers, they may become highly mobile, thus having a notable impact on possibly multiple firms’ internationalization. Exploiting a rich panel data set, the paper thoroughly tests this idea by discriminating between knowledge ascribable to managers’ former job experience and that attributable to their personal background. The paper uses a novel identification strategy grounded in on-the-job search theory to correct estimates for the presence of self-selected mobility flows. A core finding of the paper is that the maximum return to expertise acquisition is realized for those managers with previous experience in commercializing differentiated products in specific markets.

JEL-Codes: F140, F160, F230, M120.

Keywords: management, mobility, experience, export.

*Philipp Meinen\**  
*Deutsche Bundesbank*  
*Wilhelm-Epstein-Straße 14*  
*Germany – 60431 Frankfurt am Main*  
*philipp.meinen@bundesbank.de*

*Davide Sala*  
*University of Passau*  
*Dr. Hans-Kapfinger-Str. 14b*  
*Germany – 94032 Passau*  
*Davide.Sala@uni-passau.de*

*Pierpaolo Parrotta*  
*ICN Business School*  
*86 rue Sergent Blandan*  
*France – 54000 Nancy*  
*Pierpaolo.Parrotta@icn-groupe.fr*

*Erdal Yalcin*  
*University of Applied Sciences*  
*Alfred-Wachtel-Str. 8*  
*Germany – 78462 Konstanz*  
*Erdal.Yalcin@htwg-konstanz.de*

\*The paper represents the author’s personal opinions and does not necessarily reflect the views of the Deutsche Bundesbank.

June 2018

We are grateful to the Tuborg Research Centre for Globalisation and Firms and its director, Philipp Schröder, for granting us access to the data used in this paper. We are also indebted to Davide Suverato for his valuable guidance on the theoretical part and for his comments on a preliminary draft. We also benefited substantially from comments received at seminars or conferences held at the Ifo Institute, the European University Institute, the University of Mainz, the Beta Seminar, the University of Passau, University of Göttingen, and the Aarhus-Kiel Workshop. All remaining mistakes are ours.

# 1. Introduction

The international performance of companies varies greatly even in a narrowly defined industry and there is no doubt that the human factor, in the form of managerial know-how, plays a central role in explaining this empirical regularity, just as it does, in the form of managerial practices, for explaining productivity dispersion across firms (Bloom and Reenen, 2007). International business studies deem the human a unique and hard-to-imitate resource that differentiates firms in global competition.<sup>1</sup> Yet the answer to the question – to put it in the words of Syverson (2011) – “how does ‘what managers know’ affect firm performance [on international markets]?” remains elusive in the field of international economics, at least in part due to a lack of suitable data and the difficulties of establishing a causal relationship. Nevertheless, the question continues to be asked, which is not surprising given that only those firms in the highest decile of the productivity distribution – “the happy few” as labeled by Mayer and Ottaviano (2008) – account for the quasi-totality of a country’s manufacturing exports. Thus, comprehending this root mechanism is of paramount importance also for understanding national trade and its repercussions on the labor market.

Answering both the “hows” and the “whats” of this question is at the core of this paper. Framing these answers in a causal framework that builds on the economic theory grounded in the *on-the-job search* literature is the paper’s main contribution. Singling out the key attributes of managers’ export *expertise* is its other contribution.

The starting point for our argument is that catering a product to an international customer base requires knowledge that has been acquired from the practice of exporting (*learning by doing*), and as such is often tacit and informal in nature and is intimately connected to managers’ experiences and background. This makes managers, the main decision makers, a very valuable resource for a firm’s performance on international markets, but also, for the same reasons, the target of other companies’ head-hunting activity. Therefore, managers’ job-to-job mobility is an important source of knowledge transmission across firms and constitutes an important channel of how “what managers know” affects firms’ international performance: our estimates show that the increase in the probability of exporting for poaching firms can be in the order of 3-8%.

Our approach is very data intensive: it involves both the daunting task of reconstructing the mobility records of the population of managers among all firms in the economy as well as repeatedly and reliably observing firms’ international performance. To accomplish this, we exploit Danish linked employer-

---

<sup>1</sup> See Peng and York (2001) and DAILY et al. (2000).

employee registry data: the key feature of these data for our analysis is the simultaneous information on both managers' previous employers and personal background,<sup>2</sup> which allows discriminating between key attributes of managers' knowledge learned on the job and expertise derived from their *private background* (e.g., country of origin). Our analysis indicates that knowledge attributes pertaining to the *job sphere* quantitatively outclass those ascribable to managers' *private sphere*. Moreover, we find that the value of *market* expertise matured along the job career hinges on *product characteristics*, too. This is because differentiated products, contrary to homogeneous products, are ill-suited to be traded on organized exchanges or through general importers, making market expertise even more valuable when catering differentiated products on international markets.

Firms that are both exporting and poaching are hardly a random selection of the population of firms, of course, nor is their selection of managers. In a nutshell, the empirical conundrum is not to attribute varying performances on international markets to the acquisition of managerial expertise when in fact it just reflects differences in firms' unobserved base characteristics (i.e., *selection on untreated outcome*). To account for the selection of internationalizing firms that depend on poaching to sustain their internationalization strategy, the innovation in our approach is the utilization of labor theory to determine the factors on which unequal chances of hiring hinge and how to correct for these in a linear regression framework. This approach is in the same spirit as the one followed in [Dahl \(2002\)](#): like the Roy model in his approach, which allows factoring in self-selected migration in the estimation of the return to college, the [Burdett and Mortensen \(1998\)](#) model adopted here enables appropriately correcting the estimates of the firm "return" to *market expertise* acquisition in the presence of self-selected job-to-job flows.

The advantages of this approach are multiple. It relies on theoretically grounded proxies, avoiding by construction the issue of poor proxies. It is simple, as it only requires adding a correction term to the main outcome equation. It can resort to a large battery of fixed effects to filter out a great deal of unobserved heterogeneity. Finally, it avoids the quest for suitable instrumental variables, which is currently the bottleneck in the literature. Topic-wise, our paper is in the same spirit of [Labanca et al. \(2014\)](#), [Masso et al. \(2015\)](#), and [Mion et al. \(2017\)](#) as it aims to deepen our knowledge of the internationalization of firms. Methodologically, however, it substantially departs from this work. For example, it completely abandons the IV approach in favor of a proxy-based approach grounded in theory. Additionally, it adds nuance to the "what managers know" part by focusing on the interplay

---

<sup>2</sup> These data are similar to those used by [Friedrich \(2016\)](#) to study the functioning of internal labor markets and firms' recruiting strategies.

between market-specific and product expertise. Common to previous literature and our article is that using quasi-experimental evidence based on workers' displacement induced by firm closures or mass layoffs is ill-suited in this context. Indeed, although these events could be exploited as an exogenous backdrop to workers' mobility, they do not solve the selection problem related to their possible reemployment.<sup>3</sup>

Our work complements the trade and network literature, which has emphasized that immigrants are an important resource to open trade channels, making foreign labor force a prominent driver of firm internationalization (Blonigen and Wooster, 2003; Hiller, 2013; Ottaviano et al., 2018). This is because immigrants possess specific market knowledge about the market in their country of origin. However, the specificity of their knowledge derives from personal attributes (i.e., their nationality). Without neglecting this channel, our work focuses on knowledge that has been acquired throughout a manager's work career. The key differentiating factor among managers is not necessarily their nationality, but the companies in which they built their careers. In short, the most meaningful place when it comes to experience is not where the manager lived, but where he or she worked, that is, the local labor market. This means that there is scope for labor market policies to impact countries' trade through their impact on "knowledge carriers" mobility.

Obviously, our article relates to the literature emphasizing better managerial practices as drivers of firms' heterogeneous performances: With Bloom et al. (2016), it shares the view that management is akin to a technology when it comes to firm performance. However, this strand of literature has not yet looked specifically at the link between managers' mobility and firms' international performances, which is the focus of this work.

Our paper borrows its core model from the labor literature with on-the-job search. Therefore, it is also ultimately connected to the finding in this literature that the reallocation of workers across firms is an important source of firm productivity growth (Haltiwanger et al., 2017). Our paper demonstrates that such reallocation is also export enhancing.

Finally, our work relates to a large body of literature in the field of international business studies, which has for a long time considered managers' foreign experience as an important driver of firms' internationalization processes.<sup>4</sup> However, this evidence is usually based on small samples, comprised of either the largest company listed on the stock exchange market or a sample of surveyed companies,

---

<sup>3</sup> The labor literature raises doubts on their actual exogenous nature, especially in the case of managers who are in the position of "leaving the sinking ship." See Schwerdt (2011).

<sup>4</sup> See Rugman et al. (2011) for a comprehensive review.

whereas this study is based on registry data comprising the population of Danish manufacturing firms.

The structure of our paper is as follows. Section 2 presents our empirical strategy and the [Burdett and Mortensen \(1998\)](#) model, which sits at the core of our identification strategy. Section 3 briefly describes the data and presents the relevant descriptive statistics. Section 4 presents our main results on “how” managers affect firms’ international performance. Section 5 evaluates the key attributes of managers’ knowledge, the “what” managers know that shapes the internationalization process. The last section concludes.

## 2. Empirical Strategy

The goal of our empirical strategy is to be able to make a causal interpretation of the importance of acquiring market expertise for a firm’s internationalization process. Given that managers dictate and actuate firms’ strategies, it makes sense to focus on their market specific knowledge as a driver of companies’ internationalization. Management here is broadly defined and includes not only a CEO, CFO, or top-tier manager, but also middle managers, or an executive cadre. To isolate the effect of interest, we need to observe firms seeking actively to acquire managerial expertise. With employer-employee matched data, this occurs whenever expertise is acquired externally to the firm and on the domestic labor market, to a large extent by poaching. Therefore, the relevant knowledge transmission channel analyzed here is the one triggered by worker reallocation across firms (*jobs-to-jobs mobility*). This does not mean that the firm may not be pursuing other viable options (e.g., consulting services, internal promotions, etc.) to acquire such expertise, just that the pursuit of these options cannot be as easily observed in our data, although it can be controlled for in our empirical specification below.

To investigate the return of hiring “knowledge carriers” with export experience in specific markets (*event M for mobility*) for a firm’s international performances, we estimate the following outcome equation:

$$y_{imt} = \alpha_0 + \alpha_1 M_{imt} + \phi_{im} + \phi_{mt} + \phi_{it} + \underbrace{\xi_{imt} + \epsilon_{imt}}_{e_{imt}}, \quad (1)$$

where  $i$ ,  $m$ , and  $t$  index the firm, the destination market, and the time, respectively. The export outcome,  $y$ , is either an indicator of firm  $i$  exporting to market  $m$  in year  $t$  or the company’s turnover in that market.  $M$ , our binary treatment, indicates the occurrence at firm  $i$  of a hiring event involving a “knowledge carrier”, defined as someone hired externally from a company already exporting to mar-



ket  $m$  and holding a managerial position at her previous employer. This definition of  $M$  emphasizes knowledge matured on the job (at the previous employer) as opposed to knowledge pertaining to the manager’s personal background (e.g., nationality). This is not because we think the latter unimportant; but in addition to being the focus of the trade and network literature, this sort of knowledge is quantitatively less relevant in our data (see Table 1 and Section 5). Moreover, in keeping with our aim of investigating managerial mobility as being export conducive, we are measuring  $M$  as a *flow* variable that flags one or more inward movers, all with the characteristics of “knowledge carriers”. Then  $\alpha_1$  is interpretable as the average change in the outcome variable  $y$  in response to an inward mobility event that enriches a firm’s fund of expertise.<sup>5</sup>

The  $\phi$ -terms are firm-market, market-year, and firm-year fixed effects, respectively. The inclusion of  $\phi_{it}$  is especially desirable to control for other export-conducive investments and because recent evidence has shown that firms’ poaching is procyclical (Moscarini and Postel-Vinay, 2012; Haltiwanger et al., 2017; Lise and Robin, 2017) and dependent upon industry conditions (Eisfeldt and Kuhnen, 2013). Equally important, market–time fixed effects absorb demand shocks originating in the destination markets, and firm–market fixed effects control for the average stock of market specific knowledge present within the firm.<sup>6</sup>

In spite of the demanding specification comprising a complete battery of fixed effects, we nevertheless model the possibility that the treatment effect is endogenous by assuming that only  $\epsilon_{imt}$  is a truly white-noise error. Therefore, omitting the unobserved factor  $\xi_{imt}$ , which is potentially correlated with the treatment,  $\text{corr}(M_{imt}, \xi_{imt}) \neq 0$ , yields OLS estimates of the treatment effect that are biased and inconsistent.<sup>7</sup> This approach highlights that idiosyncratic and unobserved employer characteristics are problematic in this context insofar as they are market specific and time-varying. The key contribution of our identification strategy is to use, in the next section, a workhorse model from the labor literature

<sup>5</sup> Since the hiring event,  $M$  can entail one or more new hires, it is incorrect to read  $\alpha_1$  as the marginal effect of hiring an extra worker. Therefore, we are not unrealistically imposing that such marginal effect is linear across number of hires.

<sup>6</sup> See Mion et al. (2017) for an analysis measuring  $M$  as a *stock* variable. When  $M$  is measured as a stock variable, the identifying variation in the regression framework (1) comes from the yearly change in a firm’s stock of employees with market expertise  $m$ , which just reflects the difference between the “gross inflows” of workers (hires) and the “gross outflows” of workers (separations) with such expertise within the same firm. Therefore, the “stock” approach considers the net effect of both hires and separations, but, doing so, it implicitly assumes that they impact firms’ knowledge capital equally. On the contrary, our “gross inflow” based approach acknowledges that the gross outflow of workers may impact firms’ knowledge capital differently, but it neglects the effects of separations. Nevertheless, this seems appropriate in this context because it may very well be that leavers’ knowledge persists within the company even after their physical separation. The “flow” and the “stock” approaches are equally valid, but they yield similar results only in absence of separation, because only in this instance is the change in the stock of employees solely driven by “gross inflows”, ensuring that both approaches use similar variation in the data.

<sup>7</sup> The reasons for the endogenous treatment are multiple, but they can all be connected to the fact that the pool of firms poaching is not a random sample of the population of firms. With self-selection into treatment, firms may systematically differ in some base unobserved characteristics.

to proxy for  $\xi_{imt}$  and correct the estimation bias. Let  $\hat{\gamma}_{imt}$  be such a model-based proxy for  $\xi_{imt}$ : it is a good proxy if in

$$\xi_{imt} = \theta_0 + \theta_1 \hat{\gamma}_{imt} + r_{imt} \quad (2)$$

$\theta_1 \neq 0$  and, more importantly,  $r$ , the unexplained part of  $\xi$ , is “enough” white-noise that

$$\text{corr}(M_{imt}, r_{imt}) = 0. \quad (3)$$

This is the crucial identification assumption maintained throughout the paper.<sup>8</sup> Clearly, the fact that our proxy is grounded in theory helps with the requirement that  $\theta_1 \neq 0$ . But it also strengthens the untestable assumption (3), because the theory, as will become clear below, justifies modeling our proxy variable as a fixed effect, potentially capturing all sorts of firm-market-time unobserved heterogeneity.

## 2.1. Beyond Correlation: The Contribution of On-the-Job-Search Models

One problem is that the troublesome term  $\xi_{imt}$  is typically not observed, but omitting it leads to biased and inconsistent OLS estimates of the treatment effect, even controlling for a large battery of fixed effects.

Given the impossibility of simply accounting for  $\xi_{imt}$  with a fixed effect term, we use the [Burdett and Mortensen \(1998\)](#) model, a workhorse model from the *on-the-job search* literature, to proxy for it. The idea behind our approach is to exploit the model’s rich description of the job transitions occurring in the labor market to correct our outcome equation for self-selected inward mobility. Before illustrating how such a model is helpful for this purpose, this section briefly describes the essentials of the model.<sup>9</sup>

There are three reasons for choosing this model. First and foremost, it allows for *on-the-job-search*, which largely characterizes the mobility of *knowledge carriers*, typically in the form of *job-to-job* transitions. Second, it is well established in the labor literature and is analytically tractable, allowing to compute the probability that a firm is treated. Third, it exhibits wage dispersion in equilibrium (because firms compete for workers), which will be crucial for the construction of our proxy.

Let  $E_{mwt}$  be the share of workers (out of the total labor force) with expertise in market  $m$  (i.e., employed in firms exporting to market  $m$ ) and paid no more than wage  $w$  in period of length  $t$ . *On-the-*

<sup>8</sup> Note that substituting for  $\xi_{imt}$  in Equation (1) allows correctly identifying the treatment effect under assumption (3). This assumption is not testable and is the analogue of the *exclusion restriction* in the IV-approach.

<sup>9</sup> Our exposition and notation closely follows the model as presented in [Mortensen \(2003\)](#).

*job-search* means that employees search for new employment opportunities while they are employed. Therefore, their outside option is simply the job continuation value. Clearly, they would be willing to quit their jobs only if they had an offer that outclasses their job continuation value. However, because of search frictions, they have access to only a limited number of job offers, whose arrival follows a Poisson process. Using these distributional assumptions, the probability that they receive a wage offer topping their current salary and quit their current job is  $Q_{mwt}$ .<sup>10</sup>

Since  $E_{mwt}$  defines the share of potential *knowledge carriers* with expertise in market  $m$  in the entire labor force, the product of  $Q_{mwt}$  with  $E_{mwt}$  results in the share of *movers* with expertise  $m$  and who are paid no more than  $w$  in the economy. To know the fraction of them that will be hired by firm  $i$  we need to compute the probability that their winning offer was from firm  $i$ . This is just given by the share of wage offers from firm  $i$  in a  $m$ -type worker's portfolio of offers. Denoting such share with  $s_{imt}$ , firm  $i$ 's probability of hiring a knowledge carrier with profile  $m$  and paid at most  $w$  in period  $t$  is simply:

$$p_{imwt} = \underbrace{s_{imt}}_{\substack{\text{firm's} \\ \text{pull factor}}} \cdot \underbrace{Q_{mwt}}_{\text{quit probability}} \cdot \underbrace{E_{mwt}}_{\substack{\text{share of potential} \\ \text{knowledge carriers}}} \quad (4)$$

Out of the three terms, only  $s$  is firm-market specific and time varying, acting as a *pull factor* for *knowledge carriers* of type  $m$  toward firm  $i$ . As we show in the appendix,  $s$  increases with the number of job advertisements, which can be interpreted as the firm's *search effort*. The greater this is, the larger the pool of workers reached by the firm, and thus the larger the base of potential recruits. Moreover, the probability of filling an open vacancy increases in the firm's wage offer. It is therefore convenient to think of  $s_{imt}$  as a matching technology whose efficiency units increase with the firm's search effort or wage offer.

Importantly for our estimation, it is plausible that the efficiency of this technology is time varying and market specific because firms likely prioritize their use of scarce resources (either in terms of search or wage effort) based upon the strategic importance of their destination markets. For firms in our sample, different underlying values of  $s_{imt}$  translate into different probabilities of "being treated". This is problematic for the correct identification of  $\alpha_1$  in Equation (1) because the average of  $s_{imt}$  between treated and non-treated observations is likely to differ systematically in the absence of quasi-

<sup>10</sup>In the appendix we provide the exact algebraic expression of  $Q$ .

experimental evidence, leading to a selection on *untreated outcomes*.<sup>11</sup> Therefore, omitting  $s_{imt}$  (which is part of  $\xi_{imt}$ ) from Equation (1) yields inconsistent OLS estimates. Furthermore, our model suggests a positive omitted variable bias because the likelihood of hiring increases with  $s_{imt}$ . This is the case, for example, when firms with more efficient matching technology (i.e., higher  $s$ ) are also able to match more effectively not only on the domestic labor market, but also internationally with prospective buyers (i.e., higher  $y$ ). In the section below, we describe how to correct for this bias building on Equation (4).

As indicative evidence that self-selection into treatment is a concrete possibility, we show in Figure 1 that the distribution of wage offers among high productive firms is always to the right of the one of low productive firms.<sup>12</sup> Coupled with the known result that high productive firms export more on average, this result illustrates that high productive firms are more successful both at exporting and hiring. That the productivity gap between high and low productive firms is not well explained by their workforce qualities (Abowd et al., 2005; Lentz and Mortensen, 2010) points to the possibly prominent role played by non-observable employer idiosyncrasies, one of the more important of which, based on the model presented above, is the efficiency of matching technologies.

## 2.2. The Proxy Approach

Because  $s_{imt}$  depends on factors such as the number of job advertisements for which we lack data, it cannot simply be computed using its algebraic expression from the model (and reported in the appendix), but it must be estimated, so as to be included as a correction term in the outcome equation. The key to its estimation is that  $p$  in Equation (4) depends on wages, allowing us to exploit wage variation across firms' hires to estimate  $s_{imt}$  as firm-market-time fixed effects.<sup>13</sup> Operationally, we partition all workers (not just managers) along two dimensions: market knowledge and wage earned. We group destinations into eight markets (see below), and discretize the wage distribution into twenty quantiles, resulting in a grid of  $8 * 20 = 160$  cells.<sup>14</sup> Each cell defines a possible worker type with which a given firm could potentially match. Let  $h$  be an indicator variable for the realized matches, flagging mobility events in which at least one worker within wage quantile  $w$  and with expertise in market  $m$  moves to firm  $i$  in sample period  $t$ . Then, the empirical analogue of the theoretical

<sup>11</sup>This terminology is used in Blundell and Dias (2009).

<sup>12</sup>Wage offers are computed as in Christensen et al. (2005).

<sup>13</sup>Equation (4)'s dependence on wages arises in the model because the quit probability  $Q$  is fully characterizable only conditional on a given wage level. In fact, without knowing the wage level, the probability that at least one wage offer tops it cannot be determined.

<sup>14</sup>Quantiles of the wage distribution are calculated on a yearly basis.

treatment probability  $p$  derived in Equation (4) can be modeled in reduced form with the following linear probability model (LPM):<sup>15</sup>

$$P(h_{imwt} = 1|\phi) = \beta + \gamma \cdot \phi_{imt} + \zeta \cdot \phi_{wt}, \quad (5)$$

where the *pull factor*,  $s_{imt}$  from Equation (4), is modeled here as a vector of firm-market-time fixed effects,  $\phi_{imt}$ , with  $\gamma$  being the vector of coefficients associated to it and containing  $i \times m \times t$  elements. The vector  $\hat{\gamma}$  that can be estimated by OLS from Equation (5) is key to our approach because it contains, as its  $imt$ -th element, the model-based proxy,  $\hat{\gamma}_{imt}$ , that has to be added to outcome Equation (1) to correct the OLS bias from the omission of  $\xi_{imt}$ . The identifying variation for our proxy comes from firm  $i$ 's matches within the same  $m$  cell but across different  $w$  cells.<sup>16</sup> Therefore, to wipe out shocks idiosyncratic to particular wage-groups, Equation (5) further includes a vector of wage-time fixed effects,  $\phi_{wt}$ , as additional regressor. However, the sheer number of firm-market-time fixed effects to be estimated imposes a minimalist specification of Equation (5), whose unique *raison d'être* is to achieve the cleanest identification of the unobserved heterogeneity stemming from firms' different underlying values of  $s_{imt}$ . In fact, to keep estimation of  $\gamma$  feasible, it is necessary to estimate Equation (5) year by year.<sup>17</sup>

As our proxy relies on fixed effects, it is fairly general as to which firm-market-time unobserved heterogeneity it is accounting for, unless such heterogeneity is orthogonal to wage variation across firms' hires. Therefore, our approach will likely succeed in correcting for  $\xi_{imt}$  in our outcome equation, even if the latter possibly includes other factors beyond the pull factor,  $s_{imt}$ , described by our model.<sup>18</sup>

Moreover, to broaden the scope of our approach and strengthen our identification assumption (3),

<sup>15</sup>Note that  $M_{imt} = 1$  (i.e., the firm is treated) whenever  $h_{imwt}$  is not identically null for every possible wage class. And since  $p \cong P(M_{imt} = 1|\phi) = Pr\{\exists w : h_{imwt} \neq 0\} = a \text{ function of } P(h_{imwt} = 1|\phi)$ , we have that  $P(h_{imwt} = 1|\phi)$  can indeed be modeled (in reduced form) as  $p$  defined in Equation (4).

<sup>16</sup>Note that throughout the sample period, the unconditional probability of hiring at least one worker (of any wage quantile) with market-specific expertise is around 50%. Moreover, the unconditional probability of hiring at least two workers with different wage quantiles amounts to about 30%.

<sup>17</sup>For instance, in the year 2005, there are more than 4,000 firms for which we wish to estimate  $\gamma$ , implying the need to estimate more than 32,000 parameters in that single year (using close to 390,000 observations). This is indeed the reason to opt for a linear probability model in the first place but, even so, we have to resort to algorithms developed to estimate high dimensional fixed effects (see, e.g., [Abowd et al., 2002](#)). Note that we estimate Equation (5) by year and normalize the coefficients associated to firm-market fixed effects according to the mean effect in a given year. Another advantage of specifying a linear probability model is that it is free of distributional assumptions whose credibility would be limited when dealing with unobserved characteristics.

<sup>18</sup>For example, firm's stock of people with market specific knowledge  $m$  could also be part of  $\xi_{imt}$ . Clearly, such a stock ought to impact exporting to market  $m$ . And if it affects firm's success in hiring people with  $m$ -competences too, presumably because candidates on the job market with such expertise are attracted to (rejected by) an employer with already a large stock of such expertise, its omission will be biasing our estimates. However, our empirical proxy is general enough to encompass these dynamics.

the robustness analysis will also exploit sources of variation alternative to worker-wage variation: following Dahl (2002), workers will also be partitioned according to some of their personal and sociodemographic characteristics.

### 3. Data

Our information about firms and workers spans the period 1999-2012 and is from several sources made available by the Danish official statistical institute (Denmark Statistics): the *Integrated Database for Labor Market Research (IDA)* for individual-level data, the *GENERAL FIRMASTATISTIK*, for firm-level data, and the *Foreign Trade Statistics Register (UDENRIGSHANDELSSTATISTIKKEN)* for trade data.

Worker-level data contained in *IDA* are key to our proxy approach, as it contains information on workers' demographic characteristics, workplace, salary, and labor market status. Additionally, one can know the *degree of annual unemployment*, that is, the share of a year spent in unemployment.<sup>19</sup> Customs data are used to compute the two outcome variables: firm's export status and foreign sales in a given market and year.<sup>20</sup> Given the extraordinary number of fixed effects to be estimated in Equation (5), we must limit the number of markets in some way for our proxy approach to be feasible. Although the whole set of destinations is available, we group destinations into eight market groups: *Germany*, *Sweden*, *EFTA* (which includes Norway and Iceland), *EU-15*, *EU-27*, *NAFTA*, *Asia* (which includes Japan and China), and *Rest of the World*. These are mutually exclusive groups and are either a historically important market in Danish trading relations (e.g., *Germany* and *Scandinavian markets*) or a predominant trading block in modern world trade relations.<sup>21</sup>

Our sampling procedure is as follows. We begin by merging *GENERAL FIRMASTATISTIK* and

---

<sup>19</sup>The variable we use is *arledgr* and ranges continuously between 0 (always employed) and 1 (unemployed for the entire year). So a value of 0.25 would mean that a worker was unemployed for 25% of the year, or three months. Note, however, that the variable does not take into account either the number, or the duration of unemployment spells. So a worker with an uninterrupted unemployment spell of three months has the same *degree of annual unemployment* as a worker with three spells of unemployment of one month each.

<sup>20</sup>Trade data consist of customs data transactions by product and destination. The data come from two sources: *Extrastat* (for trade to a country outside the European Union) and *Intrastat* (for trade among EU member states). *Extrastat* export data come from custom forms and tax authorities and cover nearly all trade with non-EU countries, while *Intrastat* export data are self-reported figures by Danish firms following the EU regulation. Although only firms whose annual export value exceeds a certain threshold are obliged to report, *Intrastat* is a fairly complete source, accounting for approximately 97% of all exports from Denmark to EU countries in every year.

<sup>21</sup>Because groups are mutually exclusive, *EU-15* includes all EU-15 members except for Germany and Sweden, which are already grouped separately. Likewise, *EU-27* only includes those countries that are members of the European Union, but that are not already included in the *EU-15* group. In addition to China and Japan, *Asia* includes the tiger economies Hong Kong, Korea, Singapore, and Taiwan. *EFTA* countries are Switzerland, Norway, Iceland, and Liechtenstein.

*IDA* and keep only those individuals working for firms in the private sector and earning a positive wage.<sup>22</sup> We then merge export information from the “Foreign Trade Statistics Register” for these firms to learn about their export performances. For defining our treatment, we rely on the Danish occupational code (DISCO) present in the IDA data set.  $M$  is 1 if a firm hires at least one new employee poached from a firm exporting (in the previous year) to market  $m$ , and who, in her previous job, was employed as a manager (i.e., DISCO 1 - *upper management* - or DISCO 2 - *upper management, executives, or cadres*) and paid more than the firm’s median wage.<sup>23</sup> This implies that our treatment is first available in year 2000. Export experience matured at any firm counts for  $M$ , regardless of the firm’s industrial affiliation. However, outcome equation (1) is estimated only for manufacturing firms.<sup>24</sup> Moreover, the estimation sample is restricted to firms with at least 10 employees for which the key variables are not missing.<sup>25</sup> We end up with a sample comprising around 6,500 manufacturing firms and close to 350,000 firm-market-year observations.

### 3.1. Descriptive Statistics

Table 1 presents descriptive statistics of all variables in our regression. The mean of our main dependent variable, namely, the market-specific export dummy, indicates that the unconditional probability of exporting to a given market is 30%. When export status is aggregated across all markets, the share of firms that export to at least one destination is almost 60%, confirming that the Danish economy is a small open economy. Figure 2 presents the distribution of firms’ export status by market and confirms the well-known fact that the share of exporting firms is higher among high productive firms. More importantly, the figure shows that our market aggregation attains a good balance between having an adequate number of market groups and avoiding empty bins with just a few exporters.<sup>26</sup> Moreover, this market definition successfully balances between two conflicting objectives. On the one hand, each additional market included improves the accuracy with which we can measure managers’

<sup>22</sup>Hence, we drop firms in NACE rev. 2 sectors 84 and above from the sample since these sectors are dominated by public activities such as public administration, defense, education, and health care.

<sup>23</sup>The latter condition is imposed to minimize measurement error in the DISCO variable: salary in these managerial positions should be in the top half of a firm’s wage distribution. The following example clarifies the definition of our treatment: If firm  $i$  hires a DISCO 1-2 manager in 2005 from a firm exporting in  $t-1$  to *Germany* and *Norway*,  $M_{imt} = 1$  whenever  $m = \textit{Germany}$  or  $m = \textit{EFTA}$ , and  $t = 2005$ .

<sup>24</sup>To be precise, we consider firms in the NACE rev.2 two-digit sector codes 10 to 33. The only exception being firms in NACE sector 19 (manufacture of coke, refined petroleum products and nuclear fuel), which we drop from the sample.

<sup>25</sup>We drop observations with missing or implausible information, such as firms reporting zero or negative turnover, fixed and total assets, and number of employees. Additionally, we trim firms in the lower and upper 1% of the (labor) productivity distribution.

<sup>26</sup>Note that this would certainly not be the case if all destination markets were individually included because certain remote markets are served only by the most productive firms.

market-specific expertise, reducing possible measurement errors in  $M$  (and downward bias in our estimation). On the other hand, a higher number of markets translates into a higher number of fixed effects and worsens the chances of estimating Equation (5) successfully, which is essential to correct the omitted variable bias. Moreover, with more markets, the mass of zeros in our treatment indicator increases because the number of hires of at least one DISCO 1-2 manager from other firms is generally quite low and their exporting experience rather limited in its scope (after all, the majority of firms are exporting to just a few popular markets!). Indeed, Table 1 indicates that only 5% of firms are poaching throughout our sample and we know that 75% of hires involve just one manager with market specific experience. Hiring two such managers corresponds to the 90<sup>th</sup> percentile, hiring five of them occurs at the 99<sup>th</sup> percentile of the distribution of hires. The hiring of foreign managers is even rarer and only occurs in 0.1% of the firms. These low figures are expected from the approach followed here of measuring the treatment  $M$  as a flow variable.

It is likely that the restricted variation in our treatment variable and the demanding specification of Equation (1) impose a relatively high minimum data requirement in terms of the necessary length of our panel data. Accordingly, our sample period is the longest available period without changes to industry classifications.

## 4. Empirical Results

We first present, in Table 2, plain OLS estimates of outcome equation (1). In the specification comprised of just the treatment variable and the whole battery of fixed effects (Column 1), we find that hiring managers with market expertise boosts the presence of the firm on international markets. The effect is also highly significant at 1% level.

In Columns 2 and 3 of Table 2, we correct the OLS specification for the omitted factor,  $\xi_{imt}$ , by adding the proxy  $\hat{\gamma}_{imt}$ , obtained from the LPM in Equation (5). Throughout the paper, we compute two correction terms (to be used separately), differing only by the type of wage variation exploited in the estimation of the LPM (Equation (5)).

The first version of our proxy uses wage variation indistinctively across all hires of a firm; the second version restricts wage variation to hires of *separated* workers, defined in the model as those who separate involuntarily from their previous employer because of a separation shock. To identify them in the data we use the fact that they must undergo an unemployment spell in their job-to-job



reallocation, that is, their observed *degree of unemployment* must be positive.<sup>27</sup> Intriguingly, for our purposes, their job mobility does not reflect their own will, but is triggered by an exogenous cause. Adopting the terminology employed in the migration literature, it is useful to think of a new firm-worker match as resulting from two components, one related to the worker’s separation at will from her previous employer (i.e., *push factors*) and one related to the attractiveness of the hiring firm (i.e., *pull factors*). As *separated* workers’ mobility reflects exogenous factors, *push factors* cannot play an important role. Therefore, the proxy based on separated workers should help to identify  $\hat{\gamma}_{imt}$ , recipient firm’s *pull factors*, as neatly as possible.<sup>28</sup> There are problems with it, though: separated workers may not fall into our target group, that is, managers, possibly weakening the validity our identification assumption (3). This is why we always present all results with both types of proxy.

Two features stand out from our results. First, the effect of hiring managers with export experience remains positive and significant after adding the correction terms to the regression. Quantitatively, a 1 percentage point increase of our treatment, increases the export probability by 0.010 percentage points, or about 3% (most conservative estimates).<sup>29</sup> The knowledge effect described here is not found just for top managerial positions (e.g., CEO, CFO), but extends to a broader group of employees within the *cadre* figures. Second, both proxies are statistically significant, notwithstanding the inclusion of firm-year fixed effects, indicating they are indeed capturing relevant firm-market-time-specific unobserved heterogeneity. Their impact on the treatment effect is rather limited, implying just a slight decrease of the estimated coefficient, which is in line with the model’s prediction of a positive omitted variable bias.

In short, and in the spirit of providing an easy to remember take-away point, the reallocation of workers is not only *productivity enhancing* - as confirmed by a large body of the literature - but is also *export enhancing*! In what follows, we assess the robustness of our results.

#### 4.1. Robustness Checks

Key to our approach is that our proxies are capturing enough of the endogenous variation left in the data. If the endogenous variation unfiltered by all our fixed effects was orthogonal to the wage

<sup>27</sup>In this second version,  $h_{imwt} = 1$  in Equation (5) indicates that firm  $i$  is hiring in period  $t$  a *separated* worker of type  $m$  in wage class  $w$ .

<sup>28</sup>This argument assumes no “leaving the sinking ship behavior” (Schwerdt, 2011). Unfortunately, the *degree of unemployment* variable does not allow identifying the number of weeks before separation.

<sup>29</sup>These figures are obtained by dividing 0.01 by the mean of the export status (0.3) from Table 1. We favor this interpretation based on percentage points over “the hire of one extra manager” because our treatment actually refers to “hiring episodes” — the event of hiring at least someone with market expertise. Therefore a hiring episode may involve more than one person: as explained above, the great majority of hiring episodes involve up to five persons.

variation across firms' hires used for our proxy identification, both proxies should clearly be insignificant. Although this is not the case, it could be that our approach is not filtering out enough of the endogenous variation, failing assumption (3). This could result if it was the case that firms have different matching technologies for high- and low-skilled workers, rendering our correction terms based on either all workers or *separated* workers inadequate for *job-to-job* transitions involving managers.

In Table 3 we re-estimate our two proxies exploiting wage variation only across firms' hires of workers in the upper 10 quantiles. These workers are paid in a given year more than the median wage and are therefore - in terms of salary, wage offers, tasks performed on the job - closer to our managerial target group. So, if a firm is attracting these types of workers, it will likely be viewed as an attractive option by managers, too. The coefficient and significance level of the new proxies and of the treatment effect are remarkably similar to the ones found above, indicating that they are capturing the same endogenous variation as the baseline proxies. This inspires confidence that the efficiency of firms' matching technology does not vary across type of workers, a finding in line with much of the labor market literature (see, e.g., [Mortensen and Pissarides, 1999](#); [Albrecht and Vroman, 2002](#)).

There could be some concern that our results solely hinge on wage variation across firm hires. To confirm that our findings hold when a different source of individual-worker variation is used, we follow [Dahl \(2002\)](#)'s cell approach based on workers' socio-demographic characteristics. We partition workers into a grid based on *age* (four classes, 16-25, 26-35, 36-50, and 51-65), *gender*, and *education* (three classes, *university*, *secondary*, and *vocational*) for a total of 24 mutually exclusive cells. As above, each cell has to be interacted with eight categories of market expertise, yielding a total of 192 potential firm-worker matches. The sub-index  $w$  in Equation (5) is now reinterpreted as referring to one of the 24 socio-demographic cells and  $P(h_{imwt} = 1|\phi)$  is firm  $i$ 's probability of hiring a  $w$ -type worker with expertise  $m$  in period  $t$ . The identifying variation is clearly across cells and exploits hiring episodes that involve workers with different socio-demographic backgrounds. Given that many of these socio-demographic characteristics also explain wage levels in a Mincerian wage regression, the two sources of variation are highly correlated, although not identical. It is comforting that our results in terms of magnitude and statistical significance are confirmed both for the treatment effect and the correction terms.

The robustness checks speak to the solidity of our approach, which, however, would not mean much if our story is implausible. If it is managerial knowledge (spread through managers' mobility) that matters, we should expect that the estimated effect vanishes when non-managerial hires are considered. In the

last columns of Table 3 we show how the OLS results change as we gradually weaken the stringency of our managerial definition. As expected, the magnitude of the effect fades for *broader* (DISCO 1-3) managerial occupations (marginally significant) and, especially, turns insignificant for non-managerial positions (DISCO 4-9).

## 4.2. Margins of Trade

The analysis focuses on export initiation, partly because our prior is that the extensive margin of trade should most respond to this type of knowledge. However, our customs data also contain information on export sales per market, which we can exploit to either validate or falsify our prior, with the caveat that all results are intended for the population of exporters.<sup>30</sup> The results reported in Columns 2-4 of Table 4 actually indicate that acquiring specific market expertise can affect the intensive margin of trade, with the resulting sales being about 1.04 times higher than their pre-hiring level. This magnitude remains fairly stable in all specifications, regardless of the proxy used. The significance of the effect is also consistent, and usually not higher than the 10% level, making a definitive judgment as to the validity of our prior impossible.<sup>31</sup> Overall, the results for the intensive margin of trade are not as precisely estimated and do not inspire the same level of confidence as those obtained for the extensive margin of trade, but they are consistent in the sense that they are fairly stable across all specifications. However, based on the observation that none of our proxies is ever significant, doubts remain as to whether our proxy approach is as suitable for analyzing the intensive margin of trade as it was for measuring the extensive margin of trade. In fact, following the interpretation of  $s$  in Equation (4) as the efficiency of firms' labor matching technology, it is plausible that our correction terms are more likely to be capturing aspects related to the new consumer margin of [Arkolakis \(2010\)](#) rather than unobserved heterogeneity playing at the intensive margin.

## 5. “What Managers Know”

To answer the “whats” of [Syverson's \(2011\)](#) question and gain insight into the salient attributes of managerial knowledge, in this section we challenge our main set-up in three ways. First, we challenge the view that it is managerial experience that matters, and find that managers lacking any experience on the international market have no impact on firms' internationalization. Second, we challenge

---

<sup>30</sup> Although we can identify non-exporters as those firms having no foreign revenue, it would be a daunting task to come up with a Heckman model or a double-hurdle model for sales and export status with an adequate exclusion restriction.

<sup>31</sup> We remain agnostic about this, leaving it to the reader's interpretation.

that market-specific experience is key, and discover that general experience, albeit important, is not as relevant as market-specific experience. Third, we challenge the view that managerial knowledge matured along the career path is at least equipollent to knowledge related to managers' personal background.

Finally, we exploit the richness of our data to dig deeper into the specificity of managers' knowledge by looking at whether their specific knowledge is not just market oriented, but also product oriented.

To keep this section as focused and succinct as possible, we report only the results based on our baseline proxy.<sup>32</sup>

### 5.1. General Versus Specific Knowledge

In this section, we explore more closely whether it is truly specific market expertise that matters for firm performance on international markets. We provide both direct and indirect evidence in favor of this hypothesis.

The direct evidence is based on the comparison of specific market expertise with no expertise at all or, alternatively, general expertise. We proceed by modifying our treatment in two ways. First, we consider hires of managers without exporting experience, that is, of workers formerly holding a "DISCO 1-2" managerial position in a non-exporting firm. We call this treatment variable *hires (no experience)*. The second variant of the treatment considers hires of managers from other exporting firms, regardless of to where the firm exports. This variable is only slightly different from our main treatment variable, in that it does not exploit information on the destination markets served by the previous employer, but just on its exporting status. We call this (binary) treatment *hires (general experience)*. Note that neither of these new variables are market specific. Given the impossibility of relying on firm-year fixed effects, we additionally include the following firm time-varying controls, lagged one period: labor productivity, firm size (*total assets*), and the quality of the labor force (*the share of white collar worker*).<sup>33</sup> The results are presented in Table 5. Again, Column 1 reports the most basic specification estimated with firm controls instead of firm-year fixed effects. This table neatly illustrates that hires lacking market knowledge are hardly export conducive, and that general experience matters for export initiation, but not to the same extent as specific knowledge. In fact, general experience is statistically significant and relatively sizable if taken in isolation (Column 3), but both its significance and magnitude drastically drop when it is introduced contextually with our main

<sup>32</sup> See the appendix for the results obtained with the proxy based on *separated* workers.

<sup>33</sup> Note that we measure labor productivity as sales over employees, as in [Haltiwanger et al. \(2017\)](#).

treatment (Column 4). On the contrary, our treatment maintains just about the same magnitude and significance level as in our main analysis, in spite of the fact that the variable measuring *general experience* is based on a less stringent definition than our main treatment, so its variation in the data is necessarily larger.

The indirect evidence is based on conjecturing which type of firms should most value market expertise. It is sensible to expect that fairly homogeneous products can be more easily catered on international markets through general importers, negating the necessity of acquiring specific market expertise in the first place. Therefore, firms producing and selling internationally differentiated products should value market expertise the most. So, our estimates, which have so far neglected the type of product sold, should be understating the treatment effect compared to estimates based on a sample of firms manufacturing relatively differentiated products.

To put this conjecture to a test, we use a unique feature of our data set, which is that we have information on the manufactured products at the HTS-4 digit level for both exporters and non-exporters, albeit for a smaller subsample of manufacturing firms.<sup>34</sup> To measure the degree of product differentiation, we use the elasticity of substitution across varieties estimated by [Broda and Weinstein \(2006\)](#) at the same product aggregation.<sup>35</sup> For mono-product firms, there is a unique elasticity value. For multi-product firms, we simply take the weighted average of the elasticity associated with their manufactured products, where the weights correspond to the product-level sales out of total firm sales. The latter is computed throughout the sample years as total sales by four-digit product. At this firm-level aggregation, a firm will be classified as producing a *differentiated* product if the value of its weighted average elasticity is above the median elasticity value of all firms. Additionally, we compute a sector-specific indicator of product differentiation that varies across four-digit industries. This is obtained by averaging the product-level elasticity across firms within the same four-digit (NACE rev. 2) industry sector (weighted by firms' revenue market share in the industry). The median value of these industry-specific indicators is our reference threshold: Firms within an industry with a specific indicator above the industry reference threshold are classified as belonging to a *differentiated* industry.

Table 6 (Columns 1-4) presents the same specification as in the main analysis but splitting the

---

<sup>34</sup>This information is available from the "VARS" data. We opt for the four-digit product classification because it is the most disaggregated level that is also fairly consistent across time. Overall, firms for which this information is available account for 90% of the turnover in the manufacturing sector.

<sup>35</sup>Notice that the elasticity is available at the 10-digit HTS product level, where the first six digits correspond to HS product codes. We aggregate the elasticity at the four-digit HS product level by computing their median values across 10-digit codes. While the elasticity estimates are derived from U.S. data, they should provide a good approximation for the degree of product differentiation at a detailed product level.

sample between firms producing a differentiated product and a homogeneous product and between firms in differentiated industries or in homogeneous industries. The table makes clear that the results found in our main analysis are driven by firms in a differentiated industry or firms manufacturing differentiated products. For these firms, which supposedly benefit the most from market expertise, the treatment effect can double compared to the one found for the whole sample, providing indirect support for the thesis that it is specific knowledge that matters. Specifically, hires of manager(s) with export experience increases the probability of exporting by almost 2 percentage points in Column 2 (or about 5% when evaluated against the unconditional probability of exporting of 39% found in the sample of firms manufacturing differentiated products).

## 5.2. Market Expertise and Product Expertise

Overall, these results clearly indicate that the knowledge effect for the poaching firm is not only sizable, but also closely intertwined with its product characteristics. However, product differentiation at the source firm could be equally important. After all, managers carry over to the poaching firm the store of knowledge matured at their previous employer, including experience in catering differentiated products to an international customer base. To explore this possibility, we modify our treatment to flag those situations in which a firm hires a manager who worked for a firm that was not only exporting to a given market (as in our treatment), but also producing a differentiated product. In Table 6 call this new treatment variable *hires (market+product expertise)*. Relative to our main treatment, the new treatment conditions additionally for the “product differentiation” *status* of the previous employer. If this status is relevant for poaching firms’ export performance, we would expect to find an effect of even higher magnitude than implied by our main treatment. And this is, indeed, what Column 5 reveals, with the new treatment effect corresponding to an increase in the probability of exporting of almost 8%. These results suggest that managers’ exposure to given product characteristics enriches their market knowledge, further increasing their market value. This is not surprising if one thinks of managers’ expertise along two dimensions: *market expertise* and *product expertise*. The former is knowledge about a market, including its culture and best practices to penetrate it; the latter involves competence in the promotion and commercialization of differentiated products through appropriate sale channels. Only managers working for firms producing differentiated products and selling them internationally develop both types of expertise. When firms can successfully target this type of manager, the return of hiring should be higher. To test this hypothesis, we specify in Column 6 of Table 6 the fully saturated

model.<sup>36</sup> The positive and statistically significant coefficient on *hires (market+product expertise)* confirms that the impact of hiring a manager with market expertise increases if the manager also has *product* expertise. The return of acquiring market expertise remains statistically significant and in the range of 0.02 percentage points (or an 8% increase in export probability), but more than 75% of it is realized only when the poached manager also has the product expertise to cater differentiated goods on the international market.<sup>37</sup>

A tentative economic explanation for this is that acquiring expertise from firms manufacturing differentiated products increases also the likelihood of hiring a manager possessing both type of expertise. So product characteristics, which are arguably easier to observe than managerial expertise, function in this context as a signal for poaching firms. However, to lure managers, poaching firms have to bid up their wage offers. Firms that have an interest in doing so are just those manufacturing differentiated products because, as explained above, firms manufacturing homogeneous products cannot capitalize on product expertise. Figure 3 appears to confirm (or at least does not contradict) this interpretation: plotting again the distribution of wage offers, this time for homogeneous and differentiated product firms, shows that firms producing differentiated products tend to offer higher wage offers. They can afford to do this in the expectation of higher returns to their human capital investment, expectations that, according to our estimations, are subsequently met.

### 5.3. Knowledge and Networks

To this point, our focus has been on knowledge that managers have accumulated during their careers at competing firms, ignoring other forms of knowledge. For instance, the trade network literature finds managers' origins to be important because foreign people, due to their personal background, may have connections to the relevant consumer base. Against this backdrop, we add to our OLS corrected specification the variable *hires (foreign)*, indicating whether the poached manager is originally coming from a country belonging to market group *m*. This variable is firm-market-time specific, so its effect could be confounded with our treatment effect if it is not properly factored out by our correction terms. In Column 7 of Table 6 we show that this variable is never significant when included along with our main treatment variable. Likewise, from the interaction term included in the fully saturated model of

---

<sup>36</sup> Think of *hires (market+product expertise)* as the interaction term between our treatment variable *hires (market expertise)* and a second variable *hires (product expertise)*, which is not market specific and is therefore absorbed by firm-year fixed effects.

<sup>37</sup> The F-test for the joint significance of *hires (market expertise)* and *hires (market+product expertise)* has a p-value of 0.005.

Column 8 we deduce that our treatment effect is not noticeably different when hired managers have a foreign background. However, these results should be interpreted cautiously, as they are based on rare events in our data (i.e., the unconditional probability of hiring foreign managers reported in Table 1 is below 0.1%). Nevertheless, they do constitute sufficient evidence that tacit knowledge acquired on the domestic labor market and perpetuated through job-to-job mobility is a crucial knowledge transmission mechanism across firms that can enhance firms' internationalization processes regardless of knowledge stemming from managers' origins. In fact, our treatment effect results are robustly positive and highly significant in all specifications.



## 6. Conclusion

The evidence presented in this paper demonstrates that managerial mobility is, quantitatively, a non-negligible driver of firm internationalization. In a way, this is just further confirmation that the reallocation of workers is important for firm performance. In other ways, however, this is an important finding because it highlights the fact that the international dimension of firms does not hinge only on technological investments or internal resources, but also on labor market dynamics. This opens up a new role for labor market policy, and also suggests that the “search behavior among employed” and “who moves up the job ladder” are intimately connected with firms’ international dimension. Moreover, [Haltiwanger et al. \(2017\)](#) as well as [Suverato \(2014\)](#) and [Friedrich \(2016\)](#) emphasize the complex selection process underlying workers’ reallocation: correcting our estimates for selected mobility flows is at the core of the novel identification strategy introduced in the paper. The strategy is inspired by the approach implemented in [Dahl \(2002\)](#): the *on-the-job-search* workhorse model of [Burdett and Mortensen \(1998\)](#) used here plays the same role as his Roy model in correcting OLS bias and inconsistent estimates. The empirical results are in line with the underlying theory. Indeed, the introduced correction terms successfully capture the unobserved heterogeneity pointed out by our model. Furthermore, they tend to reduce the size of the treatment effect, consistent with the model’s prediction of a positive omitted variable bias.

Finally, the paper investigates what type of managerial knowledge makes the acquisition of external managers export conducive. Foremost, our results indicate that the experience garnered during a manager’s career, and especially that derived at the former employer on the domestic labor market, is at least as important as experience pertaining to managers’ personal background. Moreover, it is market-specific expertise that matters, especially in conjunction with the degree of product differentiation. In fact, we document the interplay between experience and product characteristics, in the sense that the maximum return to expertise acquisition is realized for those managers with previous experience at commercializing differentiated products in specific markets.

## References

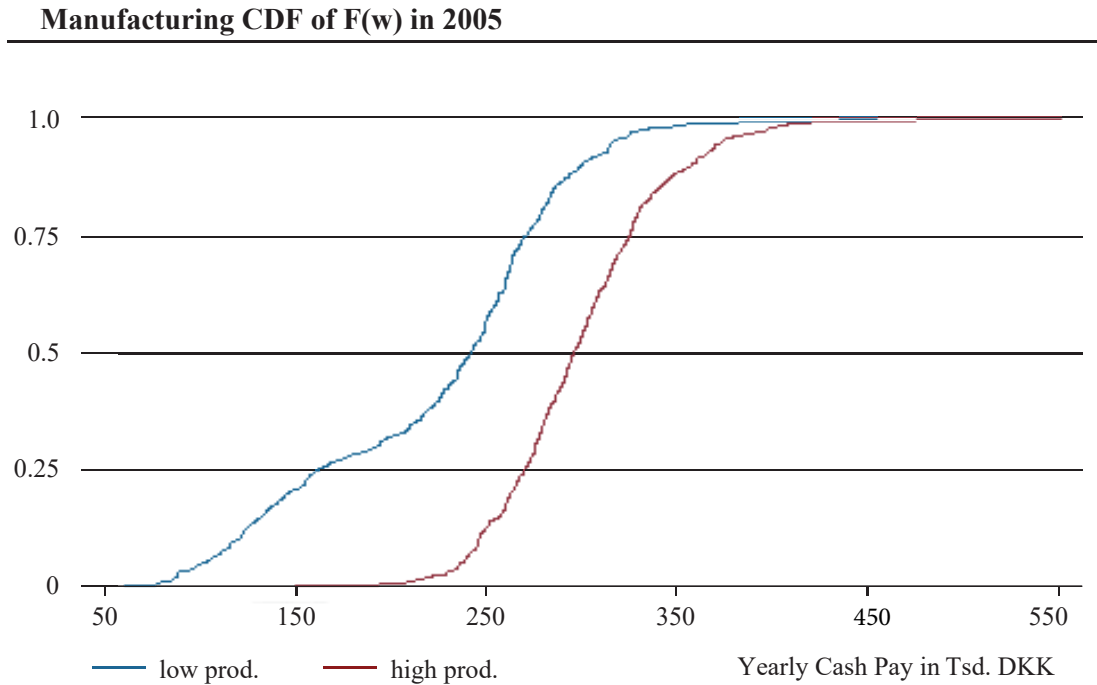
- Abowd, J. M., Creecy, R. H., and Kramarz, F. (2002). Computing person and firm effects using linked longitudinal employer-employee data. Longitudinal Employer-Household Dynamics Technical Papers 2002-06, Center for Economic Studies, U.S. Census Bureau.
- Abowd, J. M., Haltiwanger, J., Jarmin, R., Lane, J., Lengermann, P., McCue, K., McKinney, K., and Sandusky, K. (2005). The Relation among Human Capital, Productivity, and Market Value: Building Up from Micro Evidence. In Measuring Capital in the New Economy, 65, Studies in Income and Wealth, ed. Carol Corrado, John Haltiwanger, and Daniel Sichel, pages 153–204. University of Chicago Press.
- Albrecht, J. and Vroman, S. (2002). A matching model with endogenous skill requirements. International Economic Review, 43(1):283–305.
- Arkolakis, C. (2010). Market Penetration Costs and the New Consumers Margin in International Trade. Journal of Political Economy, 118(6):1151–1199.
- Blonigen, B. and Wooster, R. B. (2003). Ceo turnover and foreign market participation. Working Paper 9527, National Bureau of Economic Research.
- Bloom, N. and Reenen, J. V. (2007). Measuring and explaining management practices across firms and countries. The Quarterly Journal of Economics, 122(4):1351–1408.
- Bloom, N., Sadun, R., and Reenen, J. V. (2016). Management as a technology? Working Paper 22327, National Bureau of Economic Research.
- Blundell, R. and Dias, M. C. (2009). Alternative approaches to evaluation in empirical microeconomics. Journal of Human Resources, 44(3).
- Broda, C. and Weinstein, D. E. (2006). Globalization and the gains from variety\*. The Quarterly Journal of Economics, 121(2):541–585.
- Burdett, K. and Mortensen, D. T. (1998). Wage differentials, employer size, and unemployment. International Economic Review, 39(2):257–273.
- Cahuc, P., Postel-Vinay, F., and Robin, J.-M. (2006). Wage bargaining with on-the-job search: Theory and evidence. Econometrica, 74(2):323–364.

- Christensen, B. J., Lentz, R., Mortensen, D. T., Neumann, G. R., and Werwatz, A. (2005). On-the-job search and the wage distribution. Journal of Labor Economics, 23(1):31–58.
- Dahl, G. B. (2002). Mobility and the return to education: Testing a roy model with multiple markets. Econometrica, 70(6):2367–2420.
- Daily, C. M., Certo, S. T., and Dalton, D. R. (2000). International experience in the executive suite: the path to prosperity? Strategic Management Journal, 21(4):515–523.
- Eisfeldt, A. L. and Kuhnen, C. M. (2013). Ceo turnover in a competitive assignment framework. Journal of Financial Economics, 109(2):351 – 372.
- Friedrich, B. (2016). Internal labor markets and the competition for managerial talent.
- Haltiwanger, J., Hyatt, H., and McEntarfer, E. (2017). Who moves up the job ladder? Working Paper 23693, National Bureau of Economic Research.
- Hiller, S. (2013). Does immigrant employment matter for export sales? Evidence from Denmark. Review of World Economics (Weltwirtschaftliches Archiv), 149(2):369–394.
- Labanca, C., Molina, D., and Muendler, M.-A. (2014). Preparing to export.
- Lentz, R. and Mortensen, D. T. (2010). Labor market models of worker and firm heterogeneity. Annual Review of Economics, 2(1):577–602.
- Lise, J. and Robin, J.-M. (2017). The macrodynamics of sorting between workers and firms. American Economic Review, 107(4):1104–35.
- Masso, J., Rõigas, K., and Vahter, P. (2015). Foreign market experience, learning by hiring and firm export performance. Review of World Economics (Weltwirtschaftliches Archiv), 151(4):659–686.
- Mayer, T. and Ottaviano, G. I. P. (2008). The happy few: The internationalisation of european firms. Intereconomics, 43(3):135–148.
- Mion, G., Opromolla, L. D., and Sforza, A. (2017). The diffusion of knowledge via managers’ mobility. GEE Papers 0087, Gabinete de Estratégia e Estudos, Ministério da Economia.
- Mortensen, D. and Pissarides, C. (1999). New developments in models of search in the labor market. In Ashenfelter, O. and Card, D., editors, Handbook of Labor Economics, volume 3, Part B, chapter 39, pages 2567–2627. Elsevier, 1 edition.

- Mortensen, D. T. (2003). Wage Dispersion - Why Are Similar Workers Paid Differently? The MIT Press.
- Moscarini, G. and Postel-Vinay, F. (2012). The contribution of large and small employers to job creation in times of high and low unemployment. American Economic Review, 102(6):2509–39.
- Ottaviano, G. I., Peri, G., and Wright, G. C. (2018). Immigration, trade and productivity in services: Evidence from u.k. firms. Journal of International Economics, pages –.
- Peng, M. W. and York, A. S. (2001). Behind intermediary performance in export trade: Transactions, agents, and resources. Journal of International Business Studies, 32(2):327–346.
- Rugman, A., Verbeke, A., and Nguyen, Q. T. K. (2011). Fifty years of international business theory and beyond. Management International Review, 51(6):755–786.
- Schwerdt, G. (2011). Labor turnover before plant closure: Leaving the sinking ship; vs. &quot;Captain throwing ballast overboard&quot;; Labour Economics, 18(1):93–101.
- Suverato, D. (2014). Export and the labor market: a dynamic model with on-the-job search.
- Syverson, C. (2011). What determines productivity? Journal of Economic Literature, 49(2):326–65.

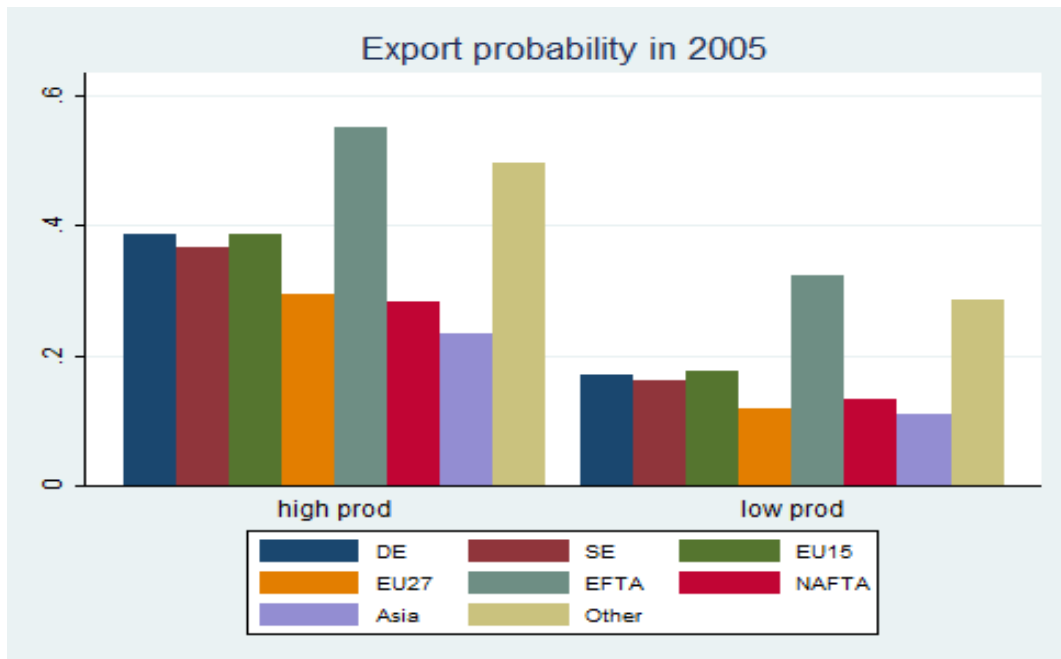
## Figures

**Figure 1:** Distribution of wage offers and export intensity by labor productivity\*

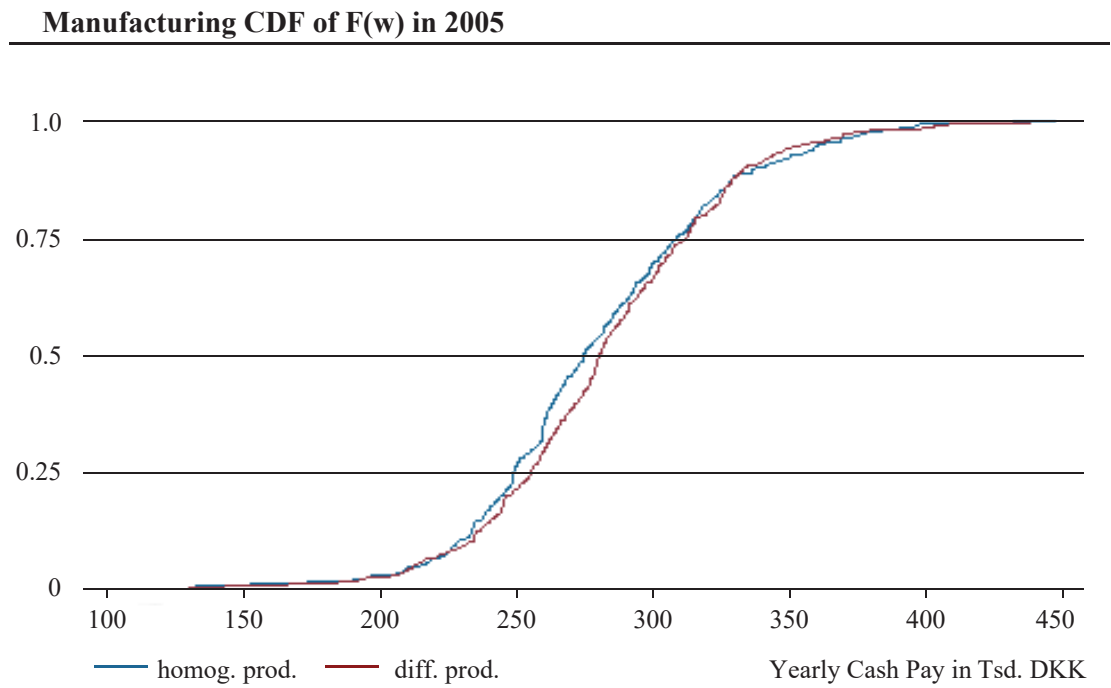


\*The computation of wage offers follows [Christensen et al. \(2005\)](#)

**Figure 2:** Share of exporting firms by productivity and destinations



**Figure 3:** Distribution of wage offers by product-type\*



\*The computation of wage offers follows [Christensen et al. \(2005\)](#)

## Tables

**Table 1:** Descriptive Statistics

Variable	Mean	Std. dev.
Export Status (market specific)	0.301	0.459
Export Status (all destinations)	0.594	0.491
Hires - <i>market expertise (M)</i>	0.053	0.224
Hires - <i>no experience</i>		
Hires - <i>general experience</i>		
Hires - <i>foreign</i>	0.001	0.028
Hires - <i>broad managers</i>	0.101	0.302
Hires - <i>non-managerial</i>	0.276	0.447
Log Labor Productivity (t-1)	13.755	0.589
Log Total Assets (t-1)	16.507	1.341
Share white-collar (t-1)	0.098	0.112
Observations	349,264	

*Note:* Observations are across firms (6490), markets (8), years (2000-2012). Markets are Germany, Sweden, EFTA, EU-15, EU-27, NAFTA, Asia, RoW.

**Table 2:** Main Results (Proxy Approach)

Dependent variable: export status	OLS		
	1	2	3
Hires - <i>market expertise (M)</i>	0.014*** (0.005)	0.010** (0.005)	0.013*** (0.005)
proxy ( <i>all workers</i> )		0.039*** (0.012)	
proxy ( <i>separated workers</i> )			0.038* (0.020)
Firm-market FE	✓	✓	✓
Market-time FE	✓	✓	✓
Firm-time FE	✓	✓	✓
R-squared	0.836	0.836	0.836
Observations	349,264		

*Note:* Hires - *market expertise* are hires of “knowledge carriers” coming from companies exporting to specific markets and who in their former job were managers (DISCO code starting with 1 or 2). Proxies are estimated with the linear probability model given in Equation (5). *Separated workers* are separating involuntarily from their employer. Standard errors are reported in parentheses. \*, \*\*, \*\*\* indicate statistical significance at 10%, 5%, and 1% level, respectively.

**Table 3:** Robustness Checks

Dependent variable: export status	OLS					
	1	2	3	4	5	6
Hires - <i>market expertise</i> ( $M$ )	0.011** (0.005)	0.013*** (0.005)	0.011** (0.005)	0.013** (0.005)		
Hires - <i>broad managers</i>					0.006 (0.004)	
Hires - <i>non-managerial</i>						-0.003 (0.003)
proxy ( <i>all workers</i> - high wage)	0.022** (0.009)					
proxy ( <i>separated workers</i> - high wage)		0.034** (0.017)				
proxy ( <i>all workers</i> - sociodemographic)			0.032*** (0.010)			
proxy ( <i>separated workers</i> - sociodemographic)				0.028* (0.017)		
proxy ( <i>all workers</i> )					0.038*** (0.012)	0.051*** (0.013)
Firm-market FE	✓	✓	✓	✓	✓	✓
Market-time FE	✓	✓	✓	✓	✓	✓
Firm-time FE	✓	✓	✓	✓	✓	✓
R-squared	0.836	0.836	0.836	0.836	0.836	0.836
Observations	349,264					

*Note:* Hires - *market expertise* are hires of “knowledge carriers” coming from companies exporting to specific markets and who in their former job were managers (DISCO code starting with 1 or 2). Ditto for Hires - *broad managers* (DISCO code starting with either a 1, 2, or 3) and for Hires - *non-managerial* (DISCO code starting with a number between 4 and 9). Proxies are estimated with the linear probability model given in Equation (5). *Separated workers* are separating involuntarily from their employer. *High-wage workers* are earning a salary above the median wage. Sociodemographic characteristics are *age*, *gender*, and *education*. Standard errors are reported in parentheses. \*, \*\*, \*\*\* indicate statistical significance at 10%, 5%, and 1% level, respectively.



**Table 4:** The Intensive Margin of Trade

Dependent variable: export turnover (log)	OLS				
	1	2	3	4	5
Hires - <i>market expertise</i> ( $M$ )	0.040** (0.020)	0.037* (0.020)	0.037* (0.020)	0.036* (0.020)	0.037* (0.020)
proxy ( <i>all</i> workers)		0.064 (0.122)			
proxy ( <i>separated</i> workers)			0.270 (0.203)		
proxy ( <i>all</i> workers - high wage)				0.061 (0.090)	
proxy ( <i>all</i> workers - sociodemographic)					0.068 (0.115)
Firm-market FE	✓	✓	✓	✓	✓
Market-time FE	✓	✓	✓	✓	✓
Firm-time FE	✓	✓	✓	✓	✓
R-squared	0.868	0.868	0.868	0.867	0.868
Observations	95,631	95,631	95,631	95,515	95,515

*Note:* Hires - *market expertise* are hires of “knowledge carriers” coming from companies exporting to specific markets and who in their former job were managers (DISCO code starting with 1 or 2). Proxies are estimated with the linear probability model given in Equation (5). *Separated* workers are separating involuntarily from their employer. *High-wage* workers are earning a salary above the median wage. Sociodemographic characteristics are *age*, *gender*, and *education*. Standard errors are reported in parentheses. \*, \*\*, \*\*\* indicate statistical significance at 10%, 5%, and 1% level, respectively.

**Table 5:** Market Specific Knowledge I

Dependent variable: export status	OLS			
	1	2	3	4
Hires - <i>market expertise (M)</i>	0.013*** (0.002)			0.009** (0.004)
Hires - <i>no experience</i>		0.002 (0.003)		
Hires - <i>general experience</i>			0.011*** (0.002)	0.005* (0.003)
proxy ( <i>all workers</i> )	0.016*** (0.005)	0.022*** (0.005)	0.018*** (0.005)	0.016*** (0.005)
Firm-market FE	✓	✓	✓	✓
Market-time FE	✓	✓	✓	✓
Firm-time FE	×	×	×	×
Firm Controls	✓	✓	✓	✓
R-squared	0.786	0.786	0.786	0.786
Observations	349,264			

*Note:* Hires - *market expertise* are hires of “knowledge carriers” coming from companies exporting to specific markets and who in their former job were managers (DISCO code starting with 1 or 2). Hires - *no experience* are hires of managers from non exporting companies. Hires - *general experience* are hires of managers from other companies that are exporting to at least one destination. Proxies are estimated with the linear probability model given in Equation (5). Firm controls include the logarithm of *labor productivity*, the logarithm of *total assets*, and the *share of white-collar workers*, all lagged one period. Standard errors are reported in parentheses. \*, \*\*, \*\*\* indicate statistical significance at 10%, 5%, and 1% level, respectively.

**Table 6:** Market-Specific Knowledge II

Dependent variable: export status	OLS							
	Product Differentiation				Product Expertise		Foreign Origins	
	<i>Homogenous Firm</i>	<i>Differentiated Firm</i>	<i>Homogenous Industry</i>	<i>Differentiated Industry</i>				
	1	2	3	4	5	6	7	8
Hires - <i>market expertise (M)</i>	-0.002 (0.007)	0.019** (0.008)	0.006 (0.008)	0.015** (0.007)		0.005 (0.006)	0.010** (0.005)	0.010** (0.005)
Hires - <i>market+product expertise</i>					0.022** (0.009)	0.017* (0.010)		
Hires - <i>foreign</i>							-0.003 (0.018)	-0.006 (0.026)
<i>market expertise (M) × foreign</i>								0.006 (0.035)
proxy ( <i>all workers</i> )	0.022 (0.020)	0.079*** (0.021)	0.019 (0.016)	0.056*** (0.017)	0.040*** (0.012)	0.038*** (0.012)	0.039*** (0.012)	0.039*** (0.012)
Firm-market FE	✓	✓	✓	✓	✓	✓	✓	✓
Market-time FE	✓	✓	✓	✓	✓	✓	✓	✓
Firm-time FE	✓	✓	✓	✓	✓	✓	✓	✓
R-squared	0.825	0.824	0.839	0.834	0.836	0.836	0.836	0.836
Observations	131,472	123,424	184,936	163,048	349,264	349,264	349,264	349,264

*Note:* Hires - *market expertise* are hires of “knowledge carriers” coming from companies exporting to specific markets and who in their former job were managers (DISCO code starting with 1 or 2). Hires - *market+ product expertise* are hires of “knowledge carriers” who in their former job were DISCO 1-2 managers in companies that are exporting to specific markets AND producing differentiated products. Hires - *foreign* are hires of managers coming from a country belonging to a specific market group. Proxies are estimated with the linear probability model given in Equation (5). Standard errors are reported in parentheses. \*, \*\*, \*\*\* indicate statistical significance at 10%, 5%, and 1% level, respectively.

## Appendix

### A. Microfoundation of firm's treatment probability

The aim in this section is to show the necessary steps to derive Equation (4) of the paper. We proceed with deriving each term forming Equation (4) separately, as in the article we have already explained why multiplying them together results in  $p_{imwt}$ .

The derivation of each term closely follows the treatment in [Mortensen \(2003\)](#) and necessitates first introducing the salient assumptions underlying the model and some notation.

- The economy is composed of  $M$  identical employers and  $N$  identical workers.
- Workers and employers maximize expected (lifetime) income as opposed to maximizing utility.
- Employers have a linear technology relating the number of workers employed to output.
- Search frictions:
  - No job-seeking worker knows the wage paid by any employer at the beginning of the period.
  - Firms have a limited capacity to inform workers about wages. To search, each firm randomly mails a number of offers, which is small compared to the number of workers on the market. So each firm reaches only a subset of workers.
  - Each employer anticipates that workers can receive more than one offer and sets the wage offer taking into account that other employers will have similar incentives for wanting to attract the worker.<sup>1</sup>
  - Some workers may receive multiple offers: in this instance, they apply for the highest-paying job. If an unemployed worker decides to decline all offers received, he or she continues to search and remains unemployed. If an employed worker does not receive a better offer, he or she continues with the current match.
- The total labor force in this model is normalized to a continuum of individuals of unity mass.
- All job transition processes are Poisson:  $\lambda$  is the *arrival rate* of offers and  $\delta$  is the *job-destruction rate*.

---

<sup>1</sup> Firms pursue a *stationary* wage policy, in the sense that the wage offered and accepted continues to be the wage paid throughout the whole tenure of the employer-employee match. Therefore, wage offers in the model are independent of tenure and experience.

- $G(w)$  is the c.d.f. of the distribution of wages.
- $F(w)$  is the c.d.f. of the distribution of wage offers. The support of  $F$  ranges between the workers' identical reservation wage  $R$  and the highest wage offered  $\bar{w}$ , i.e.,  $F : [R, \bar{w}] \rightarrow [0, 1]$ .

Although the model is solved in continuous time, for our empirical purposes it is more convenient to work with a generic time interval  $[0, t]$ ,  $t \geq 0$ .

### Derivation of $E_{mwt}$

It is easy to see that the fraction of workers who earn  $w$  or less in the economy is simply  $E(w) \equiv (1 - u)G(w)$ , where  $u$  is the unemployment rate in the economy.

In the paper, we have defined “knowledge carriers”, firms' target, as those DISCO 1-2 managers employed in firms exporting to market  $m$ . Let their fraction (in total labor force) simply be  $E_{mwt}$ , which is a part of  $E(w)$ .

### Derivation of $Q_{mwt}$

In the model, the probability of quitting a job paid no more than  $w$  in a period of length  $t$  is

$$\begin{aligned}
Q(F(w), \lambda t) &= \sum_{x=0}^{\infty} [1 - F(w)^x] \frac{e^{-\lambda t} (\lambda t)^x}{x!} \\
&= e^{-\lambda t} \sum_{x=0}^{\infty} \frac{(\lambda t)^x}{x!} - \sum_{x=0}^{\infty} \frac{e^{-\lambda t} (\lambda t F(w))^x}{x!} \\
&= 1 - e^{-\lambda t [1 - F(w)]}, \tag{A.1}
\end{aligned}$$

which is based on the hypothesis that the number of offers  $X_t$  received in the interval  $[0, t]$  follows a Poisson process with parameter  $\lambda t$ , so that the probability of receiving  $x$  offers,  $Pr\{X_t = x\}$ , is just the Poisson probability density function evaluated at  $x$ .<sup>2</sup> Since a worker will quit a job with salary  $w$  only for a better-paid job, quitting will occur only if there is a job offer higher than  $w$  among the  $x$  offers received. The probability of quitting given that  $x$  offers are received is the probability that at least one offer among those received is above  $w$ , which is  $1 - F(w)^x$ .<sup>3</sup> Using this conditional probability,

<sup>2</sup> The third equality uses the result  $\sum_{x=0}^{\infty} c^x/x! = e^c$ . Note also that the quitting rate in discrete time is defined as  $Q(F(w), \lambda t)/t$ , which for an infinitesimal time interval converges to the known continuous time formulation:  $\lim_{t \rightarrow 0} Q(F(w), \lambda t)/t = \lim_{t \rightarrow 0} e^{-\lambda t [1 - F(w)]} \lambda [1 - F(w)]/1 = \lambda F(1 - w)$ , where the first equality uses l'Hôpital rule.

<sup>3</sup> Because of the randomness of offers, receiving offers from distinct employers are independent events. Therefore, given that  $F(w)$  is the probability of receiving an offer less than  $w$ , the probability of receiving  $x$  offers less than  $w$  is simply  $F(w)^x$ .

the probability of receiving  $x$  offers and thereafter quitting the job is  $[1 - F(w)^x]Pr\{X_t = x\}$ , which is exactly the term within the mathematical summation in the first line of the equation. Summing this term over the entire support of the Poisson probability density function gives the probability  $Q$  of quitting a job paid  $w$  after receiving any number of offers.

The empirical evidence shows that the parameter,  $\lambda$ , and the c.d.f. of wage offers,  $F(w)$ , can vary by industry or occupation (Christensen et al., 2005; Cahuc et al., 2006). In particular, managers exhibit the lowest offer arrival rate, consistent with the observation that they typically experience shorter unemployment spells. Without loss of generality, let then  $\lambda_m$  be the offer arrival rate, and likewise, let  $F_m(w)$  be the c.d.f of wage offer specifically among managers of type  $m$ . Then,  $Q_{mwt}$  obtained by replacing  $\lambda$  and  $F(w)$  in Equation (A.1) with  $\lambda_m$  and  $F_m(w)$  is the probability of quitting a job paid at most  $w$  in a period of length  $t$ .

### Derivation of $s_{imt}$

Whether “knowledge carriers” with expertise in market  $m$  (henceforth, type- $m$  workers) are successfully hired by firm  $i$  crucially depends on whether they have firm  $i$ ’s job offer(s) in hand when quitting their job. This happens with probability given by the proportion of offers from firm  $i$  in their typical portfolio of offers. By assumption (Poisson processes), we know they receive offers at rate  $\lambda_m$  and the average size of their portfolio in a time interval of length  $t$  is  $\lambda_m t$ .<sup>4</sup> Now, letting  $v_{imt}$  be the number of offers mailed by firm  $i$  in this time interval to fill a type- $m$  worker vacancy, the probability that  $y$  of them reach a randomly selected type- $m$  worker is equivalent to calculating the probability of  $y$  “success” in a sequence of  $v_{imt}$  Bernoulli “trials”. Such probability is described by the binomial probability density function with parameters  $v_{imt}$  and  $p = 1/N$ , where  $p$  indicates the probability of “success” in each of the trials. This distribution also implies an average number of success of  $(v_{imt})/N$ .<sup>5</sup> Therefore, the share of  $m$ -offers from firm  $i$  in a typical offer portfolio is on average  $s_{imt} \equiv (v_{imt})/N(\lambda_m t)$ .

<sup>4</sup> Recall that the random variable  $X_t$  distributed with a Poisson probability density function with parameter  $\lambda t$  has mean and variance equal to  $E(X_t) = Var(X_t) = \lambda t$ .

<sup>5</sup> Formally, letting  $Y_{n,t}$  be the number of successes in the time interval  $(0, t]$  for a sequence of  $n$  Bernoulli trials, each with probability of success of  $p$ , then  $Pr\{Y_{n,t} = y\} = \binom{nt}{y} p^y (1-p)^{(nt)-y}$ . The mean and the variance are, respectively,  $E(Y_{n,t}) = (nt)p$  and  $Var(Y_{n,t}) = (nt)p(1-p)$ . Moreover, the distribution of  $Y_{n,t}$  converges to the distribution of  $X_t$ , which is Poisson with parameter  $\lambda t$ , as  $n \rightarrow \infty$ . Given that it is unclear whether the latter condition is met in our case, using the Poisson approximation did not seem appropriate in this context.

## B. Section 5: additional results

In what follows, we report the same tables as in Section 5, with the unique difference being the proxy used (*separated* workers).

**Table B.1:** Market-Specific Knowledge I

Dependent variable: export status	OLS			
	1	2	3	4
Hires - <i>market expertise</i> ( $M$ )	0.015*** (0.002)			0.010*** (0.004)
Hires - <i>no experience</i>		0.002 (0.003)		
Hires - <i>general experience</i>			0.012*** (0.002)	0.005* (0.003)
proxy ( <i>separated</i> workers)	0.011 (0.009)	0.016* (0.009)	0.012 (0.009)	0.011 (0.009)
Firm-market FE	✓	✓	✓	✓
Market-time FE	✓	✓	✓	✓
Firm-time FE	×	×	×	×
Firm Controls	✓	✓	✓	✓
R-squared	0.786	0.786	0.786	0.786
Observations	349,264			

*Note:* Hires - *market expertise* are hires of “knowledge carriers” coming from companies exporting to specific markets and who in their former job were managers (DISCO code starting with 1 or 2). Hires - *no experience* are hires of managers from non exporting companies. Hires - *general experience* are hires of managers from other companies that are exporting to at least one destination. Proxies are estimated with the linear probability model given in Equation (5). *Separated* workers are separating involuntarily from their employer. Firm controls include the logarithm of *labor productivity*, the logarithm of *total assets*, and the *share of white collar workers*, all lagged one period. Standard errors are reported in parentheses. \*, \*\*, \*\*\* indicate statistical significance at 10%, 5%, and 1% level, respectively.

**Table B.2:** Market-Specific Knowledge II

Dependent variable: export status	OLS							
	Product Differentiation				Product Expertise		Foreign Origins	
	<i>Homogenous Firm</i>	<i>Differentiated Firm</i>	<i>Homogenous Industry</i>	<i>Differentiated Industry</i>				
	1	2	3	4	5	6	7	8
Hires - <i>market expertise (M)</i>	-0.001 (0.007)	0.025** (0.007)	0.008 (0.008)	0.018*** (0.007)		0.008 (0.006)	0.013*** (0.005)	0.013*** (0.005)
Hires - <i>market+product expertise</i>					0.025*** (0.009)	0.018* (0.010)		
Hires - <i>foreign</i>							-0.003 (0.018)	-0.006 (0.026)
<i>market expertise (M) × foreign</i>								0.006 (0.035)
proxy ( <i>separated workers</i> )	0.008 (0.032)	0.061* (0.036)	0.003 (0.029)	0.073** (0.029)	0.039* (0.020)	0.038* (0.012)	0.038* (0.020)	0.038* (0.020)
Firm-market FE	✓	✓	✓	✓	✓	✓	✓	✓
Market-time FE	✓	✓	✓	✓	✓	✓	✓	✓
Firm-time FE	✓	✓	✓	✓	✓	✓	✓	✓
R-squared	0.825	0.824	0.839	0.834	0.836	0.836	0.836	0.836
Observations	131,472	123,424	184,936	163,048	349,264	349,264	349,264	349,264

*Note:* Hires - *market expertise* are hires of “knowledge carriers” coming from companies exporting to specific markets and who in their former job were managers (DISCO code starting with 1 or 2). Hires - *market+ product expertise* are hires of “knowledge carriers” who in their former job were DISCO 1-2 managers in companies that are exporting to specific markets AND producing differentiated products. Hires - *foreign* are hires of managers coming from a country belonging to a specific market group. Proxies are estimated with the linear probability model given in Equation (5). *Separated workers* are separating involuntarily from their employer. Standard errors are reported in parentheses. \*, \*\*, \*\*\* indicate statistical significance at 10%, 5%, and 1% level, respectively.