I Z A Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 11998

Decomposing the Exporter Wage Gap: Selection or Differential Returns?

Jonas Ehn Bødker
Jonas Maibom
Rune Vejlin

NOVEMBER 2018

I Z A Institute
of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 11998

Decomposing the Exporter Wage Gap: Selection or Differential Returns?

Jonas Ehn Bødker
Aarhus University

Jonas Maibom
Aarhus University

Rune Vejlin
Aarhus University and IZA

NOVEMBER 2018

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Decomposing the Exporter Wage Gap: Selection or Differential Returns?*

There is a large literature documenting that workers in exporting firms receive higher wages on average than workers in non-exporting firms. This is also the case for Denmark, where the unconditional exporter wage gap is 3 percent. However, little is known about the sources behind the gap: Is it because more productive (and/or higher paying) firms export, because more productive workers select into the export sector, or is it because matches in the export sector are more productive? In this paper we decompose the gap in wages into these different sources and assess their relative importance. We are the first to show that the presence of exporter-specific worker traits, that are unobservable to the econometrician, is the primary driver of the gap. To reach this finding, we employ a novel econometric strategy and exploit two state-of-art estimators. We start out with an AKM-style wage equation with worker, firm, and match fixed effects. We then use the model in a series of classical decompositions of the exporter wage gap. We show that allowing workers to have time-invariant traits specific to the exporting sector is very important for correctly assessing which factors drive the exporter wage gap. Our results suggest that workers in exporting firms have e.g. skills that are particularly valuable in the exporting sector, therefore generating higher wages in that sector. We also show that workers make job transitions based on these differential returns and thereby the exporter wage gap becomes a result of workers selecting into the export or non-export sector based on their comparative advantage. Finally, we show that our findings are not changed substantially if we instead perform the analysis in a non-linear framework instead of the linear AKM-style framework.

JEL Classification: F15, J31

Keywords: exporter wage gap, AKM decomposition, match effects

Corresponding author:
Rune Vejlin
Department of Economics and Business
Aarhus University
Fuglesangs Allé 4
8210 Aarhus V
Denmark

E-mail: rvejlin@econ.au.dk

* We would like to thank Jesper Bagger, Jakob Roland Munch, Mads Hejlesen, and participants at the Nordic Data Meeting 2017, the DTMC conference 2017 and the EEA conference 2018 for constructive comments.

1 Introduction

It is a well established fact that exporting firms pay higher wages on average. Yet, very little is known about the actual differences in the composition of workers and firms between the exporting and non-exporting sectors: Are exporting firms simply more productive? Do they hire more productive workers? Or do they simply have alternative wage policies unrelated to firm and worker productivity? The drivers of the wage gap are important to identify in order to understand the consequences of export on aggregate production and wage inequality. A stylized argument in favor of exporting is that increasing export also increases wages. This suggests that exporting also benefits workers and therefore this is used as an additional argument in favor of export promoting policies.[1] But, this argument implicitly rests on the assumption that the observed differences between the exporting and non-exporting firms are driven by firm differences from which all workers can benefit. However, the differences might as well be caused by differences on the worker side with more productive workers selecting into the exporting sector. This distinction is important for the impact of export promoting policy, since whether such policies lead to wage gains for employed workers depends on whether wage differences arise due to workers or firms. As an example, take the extreme case, where the export wage premium is a result of workers with high productivity self-selecting into the export sector. Assuming there are no match or firm specific wage components, a policy designed to increase export will not increase the total wage bill and may only lead to a restructuring of the labor market. Thus, an assessment of how the wage gap arises is important for how we should think about the effect on wages of export promoting policies. The underlying nature of the wage gap should also inform policy makers about where to focus attention when discussing international trade policies. If the exporter wage gap is primarily driven by firm differences, policies promoting export or dealing with the associated increase in wage inequality should focus on for instance addressing general conditions for firms, technology and labor market matching. However, if wage differences instead arise because of workers, the policy mix should focus on e.g. upgrading existing worker skills and providing adequate education structures for the future labor market.

In this paper we distinguish between two potentially competing mechanisms that shape the exporter wage gap. First, the selection of specific workers and firms into the export sector. Second, the effect of differential returns to characteristics (observed and unobserved) leading to alternative wage policies in exporting and non-exporting firms for the same worker. The concepts are interrelated in the sense that selection could happen *because* of differential returns across sectors. However, by "the selection effect" we mean the extent to which high and low wage workers and firms select into the export and non-export sector. The

[1] See e.g. http://www.oecd.org/tad/tradeandjobs.htm

selection effect thereby refers to a global rank of workers and firms, i.e. the extent to which the wage gap is a result of e.g. the most "productive" workers or firms being exporters instead of non-exporters. The effect of differential returns instead refers to the extent to which certain characteristics are valued differently in the export and non-export sector. For instance, the return to experience could be higher in the export sector, or similarly the return to unobserved permanent characteristics, such as language or negotiation skills, may differ since these unobservables may simply be more productive in the export sector.[2] This paper explores these different mechanisms using different econometric frameworks and wage decompositions with the main results being derived using an AKM-style framework of Abowd et al. (1999) described below.

The empirical literature has primarily focused on the direct effect of being employed in an exporting firm on wages (some key references are Bernard et al. (1995) and Bernard and Jensen (1997)). The purpose of this literature has been to recover a causal effect of firm level exports on wages. For that purpose, it is important to control for observable and unobservable differences between workers and firms. In most studies the premium is found to be economically small and sometimes statistically insignificant.[3] The AKM-style framework has been used in multiple studies of how exporting affects wages, see e.g. Frías et al. (2009) and Macis and Schivardi (2016). These papers highlight the existence of two major concerns of endogeneity in the relationship between export status and wages. 1) Firms self-select into exporting. 2) Workers self-select into firms. In these studies the worker and firm effects are recovered from an AKM-style regression under the assumption of conditional exogenous mobility. The firm-average wage, the firm-average worker effects and the firm effects are then causally related to firm level exports through an exogenous change in the trade environment (such as a currency devaluation) in a before-after style setup. Our paper has a different focus. Instead of recovering a casual effect of trade on firm level wages, we decompose the unconditional wage gap using different requirements on mobility patterns of workers. We are interested in the exporter wage gap in the whole economy and thus not the wage change induced by e.g. a trade reform. That is, knowing that workers and firms act strategically, is the differential pay between exporters and non-exporters related to the workers, the firms or the specific matches?

[2] We do not investigate *why* certain firm and workers are high or low wage entities but simply show how these differences exist in the data along different dimensions of observables. We sometimes interpret these differences as related to differences in e.g. productivity (as is standard in the theoretical work). Furthermore, we will not make any attempt to distinguish between different components of worker unobservables, but only identify a time-invariant composite index as is standard in AKM type models (See Abowd et al. (1999)). Our analysis will focus on to what extent these worker unobservables in wage equations differ between exporting and non-exporting firms. These worker unobservables could then differ because of different returns to the same worker unobservables (e.g. negotiation skills simply more valued in the export sector) or because some skills are used more intensively (e.g. language skills in exporting).

[3] In their original work Bernard et al. (1995) found that including a firm fixed effect reduced the export wage premium from roughly 10 percent to roughly 2 percent. Schank et al. (2007) find an insignificant export premium in the order of 0.5-1 percent controlling for firm, worker and spell fixed effects, respectively. Munch and Skaksen (2008) find insignificant exporter dummies controlling for match fixed effects. It should be noted that some studies find larger effects, e.g. Irarrazabal et al. (2013) find a significant effect of 5 percent controlling for worker and firm composition. Hummels et al. (2014) study the causal effect of export on the intensive margin and find a substantial elasticity of export on wages of 0.05 percent. This highlights the point that exporting can have consequences outside the effect on the extensive margin (the decision to export or not) which we focus on here. Similarly, Macis and Schivardi (2016) find that an intensive margin increase in export of 15 percent (median) leads to an increase in wages of 1.05-1.30 percent using an Italian currency devaluation. They attribute the effect to both firms rent-sharing and worker composition.

Our main contribution is to bring the unconditional exporter wage premium to the center of the analysis and to show that allowing workers to have exporter-specific skills is crucial. Our analysis involves a decomposition of the complete unconditional exporter wage gap, and is designed to combine the relatively limited direct effect of exports with a substantial and economically important unconditional exporter wage gap. We find that the main driver of the wage gap is actually that workers working in the export sector have an unobserved comparative advantage in this sector. We show specifically that they self-select into the export sector based on this comparative advantage, in particular the higher the comparative advantage, the more likely it is that workers work in the export sector. Based on this, it is perhaps not surprising that the literature has found small causal effects of exporting. Imagine using a trade reform to estimate the causal effect of exporting. The local average treatment effect, which the causal analysis is designed to identify, is identified from two marginal groups: the firms that are on the margin of exporting and workers that self-selected into a firm after the trade reform, but had not done so before. Even in the case where the intensive margin of exports is used, the identifying variation comes from marginal variation in exports and the marginal changes in endogenous worker self-selection (see Krishna et al. (2014) for an analysis of this setting). Given our results, we would not expect that the workers who respond to the reform had a strong comparative advantage in the export sector, since they did not self-select into the sector before the reform. It is very likely that the causal effect is much higher for firms and workers that are always exporting and that are exporting a lot. Thus, focusing on the reform induced variation will only provide a partial characterization of the effects of exporting on the full wage distribution. Furthermore, the unconditional exporter wage gap is what is primarily highlighted in the policy debate. Our main contribution is therefore to bring this point to the center of the analysis and provide an analysis of the sources of heterogeneity that gives rise to this. We therefore view our results as complimentary to the causal effect studies - we are fundamentally interested in quantifying the different layers of heterogeneity that the previous studies using causal analysis aim to keep fixed.

The relatively small causal effect estimates are also in accordance with the theoretical literature, which rarely features a direct effect of export, but emphasizes the role of self-selection. Theoretical trade models stress the endogeneity of firms' decision to export. However, the particular mechanisms generating this heterogeneity in the export entry decision of firms vary. Some mechanisms imply that firms are fundamentally different, some imply that they employ fundamentally different workers. In the canonical Melitz (2003) model only the most productive firms will pay a fixed cost of exporting. In this model wages do not vary across firms, but wages could increase due to an opening to export as a result of general equilibrium effects. In an extension Helpman et al. (2010) introduce search frictions and screening. This implies that exporting firms are more productive, employ workers with higher match effects and pay higher wages. In an alternative setup Yeaple (2005) focuses on heterogeneous workers. Homogeneous firms select into different sectors where the returns to worker skills vary. Exporting firms are in sectors, where returns are high. Thus, exporting firms employ higher type workers and pay higher wages. Finally, in an efficiency wage model by Davis and Harrigan (2011) exporting firms gain access to new markets and become larger. This changes the trade-off between monitoring and wages to deter shirking in favor of higher wages. Clearly, the mechanisms outlined above imply that the exporter wage gap could arise from very different layers of heterogeneity or interactions. Therefore, an empirical quantification and assessment of the importance of different sources of wage heterogeneity and their interactions in relation to export will be a useful step forward in terms of discriminating between different mechanisms and theories. Our empirical model is flexible enough to encompass these sources of wage variation.

We use population-wide Danish register data and find an unconditional exporter wage gap of 3 percent, which varies along different dimensions, which we document below.[4] We allow wage differences to arise due to observed and unobserved differences between firms, workers and matches. In our analysis, we test the importance of changing the interaction between different layers of heterogeneity and export status. In particular, our findings clearly show that allowing for both export status-specific worker and firm fixed effects is crucial for an appropriate assessment of the drivers behind the wage gap. Our final analysis contrasts the results using two different econometric models: an additive and a non-additive model. The models impose different assumptions on the interaction between worker, firm and match heterogeneity and can be seen as complementary. A simple comparison of the decompositions across these models allows us to asses the importance of critical simplifying assumptions in the additive framework. The qualitative findings are the same across the two different econometric models. Thus, the main results from the linear model are strengthened.

The additive framework is based on the job spell fixed effect model of Woodcock (2008) which includes job spell fixed effects.[5] In doing so, we allow for arbitrary correlation between time-invariant factors related to the particular match and any observable characteristics. However, the job spell fixed effects model is uninformative about the nature of the composition effects, i.e. do the job spell effects come from the worker or firm side, or is it related to the particular combination (the match).[6] To make progress in this direction, we therefore impose a set of assumptions which recover, what is essentially the two-way fixed effects model

[4] An exporter wage gap is also found in other studies using Danish data. Eriksson et al. (2009) find the gap to be 6 percent in 2003. The difference is due to the definition of export and variation over time.

[5] If the firm changes export status a new job starts.

[6] The job spell fixed effects approach is a similar econometric setup as in e.g. Krishna et al. (2014) who focus on the effect of a trade reform on worker wages while controlling for job spell fixed effects. Instead, we focus on the raw exporter wage gap and subsequently decompose the match effects into worker, firm and match components. Our approach is therefore fundamentally different from Krishna et al. (2014): In Krishna et al. (2014) the job spell fixed effects primarily serve as controls for the sorting of workers into firms based on unobservable characteristics. In contrast we are inherently interested in this sorting and aim to quantify the different dimensions of it (worker, firm and match sides respectively).

of Abowd et al. (1999) (henceforth AKM) extended with match effects that are orthogonal to worker and firm fixed effects. Within this framework, we investigate the consequences for the wage decompositions of adding different layers of interactions between export status and heterogeneity: Our baseline specification allows for export-specific firm effects, but assumes that returns to observable characteristics and worker fixed effects are the same across export status. Using this specification, we show that observable differences account for 22 percent of the exporter wage gap. However, exporting firms have higher firm fixed effects accounting for 94 percent of the gap, while workers in the export sector have lower worker fixed effects accounting for -16 percent of the gap. We extend the baseline specification by allowing for export-specific returns to observable characteristics in the tradition of Oaxaca-Blinder (Oaxaca, 1973; Blinder, 1973). In the extended specification the firm effects become even more important in explaining the exporter wage gap. The results suggest that while exporting firms pay higher wages, their wages change less based on differences in workers observable characteristics. Generally, we do not find match fixed effects to be an important driver of the exporter wage gap in the additive framework. The results imply that the primary reason exporters pay higher wages is because of firm selection, i.e. exporters have higher firm fixed effects. This is in accordance with key trade models in the tradition of Melitz (2003). Finally, our key contribution is that we allow workers to have export-specific returns to unobservables, i.e. we include export-specific worker fixed effects.[7] We follow a similar procedure as in Card et al. (2016) appropriately adapted to our setting and focus. This final extension allows us to do a full Oaxaca-Blinder decomposition separating the effect of differences in characteristics and returns for both observable and unobservable characteristics, where differences in returns to unobservables arise from a difference in e.g. worker effects under export and non-export for a particular worker. The extension turns out to be very important and drastically challenges the conclusion that firm effects are driving the exporter wage gap. In fact, what looked like higher firm fixed effects in the export sector in the baseline specification is actually the result of workers in the export sector having time-invariant characteristics that on average command higher wages in exporting firms than in non-exporting firms. As a sanity check for the econometric model, we present evidence that workers with a higher return (higher worker fixed effect) in the export sector tend to work significantly more in this sector and move into this sector with a higher propensity. The results suggests that worker selection on differential returns is an important driver of the exporter wage gap. We believe, we are the first to explicitly show that

[7] Macis and Schivardi (2016) show that the value of worker unobservables (e.g. skills) changes in response to an Italian currency devaluation which increases the exporting intensity of firms. In their setup they allow worker effects to change in the year after the currency devaluation and they show that worker effects (and wages) change more in firms that export more in response to the devaluation. The change in worker fixed effects is interpreted as a market value effect (an equilibrium effect). In this interpretation export-related skills are simply more valuable and, therefore, receive a higher return. In our setup we show that the market value of worker unobservables is "always" different depending on the export status of the firm, workers thereby have comparative advantage in one sector over another and make transitions accordingly. Thereby, we provide more empirical evidence on the importance of export related skills.

this basic economic prediction is supported by the data in the context of the exporter wage gap.

The analysis using the additive model illustrates the importance of flexible heterogeneity and its interaction with export status in a linear framework. In particular, we show that treating workers and firms asymmetrically (in terms of allowing only one side to have exporter status-specific heterogeneity) implies that results should be interpreted cautiously. With that in mind, it is reasonable to ask whether our findings change if we change the assumptions about the role of match effects, which are assumed to be orthogonal to the worker and firm effects, and the linearity assumption underlying our fixed effect model. Therefore, we investigate whether the conclusions are changed if we employ an inherently different framework, which allows for more flexible interactions between heterogeneity in the observed wage and in transition patterns. The non-additive framework comes from Bonhomme et al. (2017) (henceforth BLM). We use the BLM framework where wage profiles are non-linear. Match effects are unrestricted and are the result of interactions between particular worker types and firm classes. Importantly, workers are allowed to act on the interaction effects, thus potentially rationalizing composition patterns. The cost of this flexibility is that firms and workers must be grouped into a limited set of types. From the BLM framework, we can estimate the underlying worker and firm types, which may have been "misinterpreted" in the AKM model due to a non-linear wage schedule with the potential implication that we have attributed the gap to the wrong side of the market. As with the linear model, we decompose the raw wage gap into differences in worker and firm characteristics and compare the results. In the decompositions of the non-linear wage schedule, we employ two sets of assumptions that mirror the sets of assumptions in the baseline model and the model with export-specific worker fixed effects, respectively, in order to compare our findings and thus the implications of the assumptions. The results from the BLM model are very similar to those from the linear models. However, we find some room for match effects under the assumptions corresponding to the baseline model when we allow for worker effects to change with export. These effects are re-interpreted as differences in returns to worker effects. We find this very reassuring and it strengthens our main conclusion: The selection on differential returns to unobservables of workers is an important driver of the exporter wage gap. We are the first to apply a non-linear estimator such as BLM to study the exporter wage gap. We also want to stress that our comprehensive and innovative analysis could be easily adapted to other questions, where a decomposition of the aggregate wage gap is important for a deeper understanding of the functioning of the labor market and the appropriateness of different policies such as the gender wage gap.

The paper proceeds as follows: Section 2 contains a brief introduction to the different data sources used. The section also describes the sample selection procedure and introduces key variables related to the exporter wage gap. Section 3 establishes the linear econometric model. Section 4 extends the linear model to allow for export-specific returns to unobservables for workers. Section 3 and 4 share the same fundamental linear and additive assumption between different layers of heterogeneity. In Section 5 we asses the implications of this assumption by introducing the BLM framework which, in particular, does not impose additivity in wages. Finally, Section 6 concludes. The appendix contains further details on specific derivations and extensions, we also have an online appendix as a supplement.[8]

2 Data

We use Danish matched employer-employee data from 2004 to 2010, which we merge with data on firm level exports. Sub-section 2.1 describes the data sources, while Sub-section 2.2 describes key variables in relation to identification of the empirical model.

2.1 Data Sources

Our main data sources are the IDA database, which is employer-employee data, and the UHDI database, which contains data on firm level exports. As a final data source we also use a firm accounting database (GF) to get measures of value added, which we use only for the Card et al. (2016) normalization. We present this database in the appendix. The matched employer-employee panel contains information on workers and firms each year in November. From this data source, we extract worker level information on hourly wages, age, gender, type of employment, and occupation. We also extract workplace level data on Danish 6-digit industry indicators which nest 4-digit NACE industry codes. Hourly wages and experience are calculated using mandatory pension contributions which are used to proxy hours (see Lund and Vejlin (2016)). We only focus on jobs in the private sector.

From the firm export register UHDI we extract data on firm level exports. UHDI is constructed from two registers: Intrastat and Extrastat. Intrastat was established in 1993 with the introduction of the integrated market in the European Union (EU). Data for this register is self-reported as part of firms tax statements and it covers within-EU trade. The data in Extrastat is continuously self-reported data for goods crossing the Danish borders to non-EU countries. We define our key variable export as outside-EU exports and exclude Norway from the definition of outside-EU exports. This is also the approach taken by Munch and Skaksen (2008) and secures that we can identify all firms which are exporters.[9] We do this since within-EU trade is more similar to trade between US states than ordinary cross-border trade, as both transportation costs

[8]For convenience the online appendix is available after the appendix in this version of the paper.

[9]All firms exporting outside EU are in Extrastat. In Intrastat coverage is not complete as firms with low export and/or low import from EU-countries do not have to report within-EU trade. The exact threshold is set such that reported exports account for estimated 97 percent of total within-EU exports. In 2008 the threshold was 5.2 million Danish Kroner DST (2008). According to Statistics Denmark 40.000-50.000 firms are below the threshold each year, while additional 12.000 firms are above the threshold and thus in Intrastat (Statistics Denmark (2008)).

Table 1: **Sample Overview**

	Years	# Workers	# Firms	# Worker-Years	Exporter Wage Gap
Full sample	2004-2010	1.685.179	148.868	7.355.114	0.031
VA sample	2004-2010	1.311.385	93.627	5.502.057	0.028
BLM sample	2009-2010	859.097	61.221	1.718.194	0.046

The table shows the number of workers, firms and worker-years, as well as the raw exporter wage gap for each of the three samples used in the paper.

and other costs are likely to be low within EU (product regulations are standardized and no tariffs exist). Thus, for comparability with the literature on the export premium on US data (see e.g. Bernard et al. (1995); Bernard and Jensen (1997)), the focus on outside-EU export is meaningful.

In 2004 the European Union was expanded from 14 to 25 member states and later in 2007 expanded to 27 with the addition of Bulgaria and Romania. To obtain a consistent definition of outside-EU export, we focus on export after 2004. However, we exclude Bulgaria and Romania from the outside-EU trade for the entire period. Thus, our sample period is 2004 to 2010, which is the furthest our data allows us to go.[10]

We define a firm to be an exporter in any year, where we observe that the firm is exporting.[11] We focus on the largest set of connected firms as described by Abowd et al. (2002). Two firms are directly connected if a worker moves between them. Two firms belong to the same connected set if a chain of worker moves between firms can be formed between them. The notion of connected sets is central to the identification of the model we estimate. Our final sample (referred to as the Full Sample) has 7.355.114 worker-year observations.[12] We will also be working with two other sub-samples, the VA sample and the BLM sample. Both samples consists of the set of firms for which we can obtain information on value-added through the firm statistics database GF. We will give further details about the VA and BLM sample later. Table 1 provides an overview of the different samples we will use.

2.2 Exporters and Non-Exporters

In this section, we describe some key variables in relation to our empirical model. Table 2 shows differences of selected variables between workers in exporting and non-exporting firms.

Workers in exporting firms receive 3 percent higher wages on average over the sample period. They are on average more likely to be male, older, and more experienced. The difference in terms of years of education is not large, but there are large differences in the occupations of the workers, e.g. workers in exporting

[10] For a thorough analysis of Danish exporting firms in the period 1993-2003 see Eriksson et al. (2009).

[11] We have also tried to define exporting firms as the firms, where the exports are e.g. 50 percent or more of the total revenue and non-exporters as those that are not exporting at all. Our results are qualitatively the same.

[12] 20 percent of all firms are dropped when considering the connected set. However, only 2.66 percent of all worker-years are dropped, so the firms in the connected set still account for almost all employment and aggregate export.

In the appendix we describe our very limited sample selection in more detail, see Section B.2.

Table 2: Summary Statistics

	By Worker-Year		
	Export	Non-Export	Difference**
Wage	5.35	5.32	0.03
Experience	19.61	18.15	1.46
Age	42.54	41.76	0.78
Female	0.33	0.36	-0.02
		Education	
Less than 12 years	0.22	0.20	0.02
More than 12, less than 14	0.61	0.63	-0.02
More than 14	0.17	0.17	-0.00
		Industry	
Manufactoring	0.48	0.10	0.38
Construction	0.03	0.16	-0.13
Wholesale and retail trade	0.27	0.18	0.09
Transport, storage and communication	0.10	0.12	-0.02
Real estate, renting and business	0.06	0.19	-0.13
Other Industries	0.07	0.25	-0.19
		Occupation	
Management	0.05	0.04	0.01
Professionals	0.11	0.11	-0.00
Technicians and Ass. Professionals	0.21	0.16	0.05
Clerical support	0.11	0.11	-0.00
Sales and service	0.06	0.07	-0.01
Skilled agricultural	0.00	0.01	-0.00
Crafts	0.14	0.16	-0.03
Plant and machine operators	0.16	0.06	0.10
Elementary	0.17	0.29	-0.13
Firm size	2002.89	309.46	1693.43
Firm age	26.47	16.28	10.19
Years in sample, worker	5.71	5.31	0.40
Average match duration	4.24	3.66	0.58
	By Firm-Year		
	Export	Non-Export	Difference**
Firm size	55.93	7.67	48.25
Firm age	18.16	12.09	6.06
Extra-EU export (mil.)	19.35	0.00	19.35
Years in sample, firm	5.27	5.06	0.21
		Industry	
Manufactoring	0.36	0.08	0.28
Construction	0.02	0.17	-0.16
Wholesale and retail trade	0.52	0.23	0.29
Transport, storage and communication	0.04	0.09	-0.05
Real estate, renting and business	0.05	0.17	-0.13
Other Industries	0.02	0.25	-0.23
Value-added per worker* (1000)	535.41	379.21	156.19

The table summarizes key differences between exporting and non-exporting firms and the worker employed in those firms. *Value-added data is only available for a subset of firms (see Table 1). **All differences are significant at a 1 percent level.

firms are more likely to operate plant and machinery. This is only natural, since exporters are more likely to be manufacturing firms. Workers in exporting firms tend to be in the sample longer and have longer matches. Turning to the firm side, exporting firms are larger and are in the sample more years. For the subset of firms with value-added data, we also see that they tend to be more productive. Finally, there are large differences in industries. For instance, firms in manufacturing, wholesale and retail seem to export to a higher degree. All of this point towards massive observable differences. The exporter wage gap could simply be due to these differences between exporting and non-exporting firms and their workers. Quantifying the importance of these different dimensions (observed and unobserved) is the focus of this paper. In the next section, we set up a formal empirical model to jointly investigate and quantify the importance of different sources of heterogeneity.

In Appendix B.3 we describe patterns on the extensive margin of export status, i.e. how many firms and workers change export status and in which direction. We show that roughly 2 percent of all firms (or 3 percent of all workers) change export status in each direction *each year*, hence in our full sample we have a substantial group of workers and firms that changes export status. The variation in export status of both workers and firms is central for the identification of the empirical models we estimate below, and thus the subgroup of changers across export and non-export is an important feature of our data.

3 Decomposing The Wage Gap

The additive model is estimated in two steps following Woodcock (2008). In the first step, a job-spell fixed effects model (1) is estimated, where job spell fixed effects change when firms change export status or a worker changes employer.[13]

Our empirical model is based on the following Mincerian log-wage equation

$$w_{ijt} = \mu_{ij} + x_{ijt}\beta^{h_{jt}} + \epsilon_{ijt} \tag{1}$$

In our notation w_{ijt} is the log wage of worker i in firm j at time t. μ_{ij} is a job-spell fixed effect. x_{ijt} contains observable characteristics and can be partitioned into x^F_{jt}, which is firm level time-varying characteristics, and x^W_{it}, which are worker level time-varying characteristics. $\beta^{h_{jt}}$ indicates that the returns to observables could be different with export status, where $h_{jt} = \{E, NE\}$ gives firm j's export status at time t, and E and NE denote exporter and non-exporter, respectively. In our baseline model $\beta^{NE} = \beta^E = \beta$ and returns can

[13]Note, that if a particular match between a worker and a firm is interrupted by a period of unemployment or employment in a different firm, the original match keeps the same fixed effect when it starts again. We use the terminology of job-spell fixed effect to separate it from the "match fixed effects" that we define below.

be further partitioned into returns to worker and firm observables,β_W and β_F. In the baseline specification we include year fixed effects and up to a cubic worker and firm age profiles. This combination means, that we must exclude the linear terms of the age profiles, as both worker and firm age increase mechanically within job-spells leading to collinearity with year dummies, and hence the impact of each specific element cannot be separated (the classical problem of having cohort, time and age effects, see e.g. (Card et al., 2018)). The estimated model includes worker age squared, worker age cubic, firm age squared, firm age cubic, year fixed effects and job-spell fixed effects.[14]As we show later unobservable heterogeneity is important in explaining the log-wage variation. The R^2 for a regression with year effects, industry-gender-education effects, gender-education-specific age profiles and industry-specific firm age profiles is only 0.36. The R^2 for the job-spell fixed effects regression is, in comparison, 0.93 (adjusted R^2 is 0.88). Thus, unobserved heterogeneity that can be linked to workers and firms is important.

Equation (1) is uninformative of whether any differences in job spell fixed effects for exporters and non-exporters arise due to firms, workers, or the particular match combination of the two. To address this, we decompose the job-spell fixed effects into worker, firm and match fixed effects in a second step. We treat firms as different entities, when they are exporting and when they are not. Thus, for each worker-firm pair two hypothetical job-spell fixed effects exist, one for export and one for non-export case. Formally, we have

$$\mu_{ij} = 1\{h_{jt} = E\}\mu_{ij}^{E} + 1\{h_{jt} = NE\}\mu_{ij}^{NE}$$

where $h_{jt} = \{E, NE\}$ gives firm j's export status at time t (i denotes the worker). We impose the assumption that job-spell fixed effects arise through an additive combination of worker and firm effects, and that remaining match effects are orthogonal to these. More formally we decompose $\mu_{ij}^{h_{jt}}$ into a worker, an export status-varying firm component and an orthogonal residual match effect. I.e.

$$\hat{\mu}_{ij}^{h_{jt}} = \psi_j^{h_{jt}} + \theta_i + \tau_{ijt}^{h_{jt}} \tag{2}$$

for $h_{jt} = \{E, NE\}$. The match effect is just the average of the residual $\tau_{ijt}^{h_{jt}}$ over the length of the particular (i, j)-match. We thereby essentially recover the AKM model with the addition that residual match effects may correlate with included observables and that a firm is treated as a new firm when export status changes (In Section 4 we allow workers to have export-specific fixed effects as well). Thus, the assumptions correspond to the job spell fixed effects model of Woodcock (2008) with the addition that the firm and match

[14]Excluding the linear terms implies that part of the age profile is captured by the job-spell fixed effects, while another part is captured in the quadratic part. This introduces a negative correlation between the fixed effects and the observables. To deal with this we follow the convention and recenter worker and firm age variables at the point where the profiles are flat (see also (Card et al., 2018)). For worker age this happens at age 40, while firm age profiles are flat around age 18.

effect can vary with export status.

The model is estimated under the assumption of strict exogeneity, i.e. $E(\epsilon_{ijt}|x_{it}, h_{jt}, i, j, t) = 0$, which is also referred to as exogenous mobility as in e.g. (Abowd et al., 2015). Our assumption of export status-specific firm effects implies that workers actually are allowed to move based on the export decisions of the firms, but only based on, what is captured by the firm effects. Workers are not allowed to react to particularities (i.e. draws of ϵ_{ijt}) of either current or future matches, beyond how this is related to observed characteristics.[15]

3.1 The Exporter Wage Gap

In this section, we will present our baseline decompositions of the unconditional exporter wage gap. The decompositions are an important tool in order to discriminate between different theoretical mechanisms and to guide researchers in terms of what layers of heterogeneity models should incorporate, in order to fit empirical observations. Suppose for instance, we found no difference in the firm fixed effect component. Although we do not propose a particular test, we would consider this evidence against a Melitz (2003) type mechanism. In that model the exporter wage gap would result from intrinsically higher productivity of the firms self-selecting into exporting. If firm differences in wages are not particularly important in the data this suggests that the role of firms in generating the exporter wage gap is less important.

3.1.1 Decomposition with Homogeneous Returns

First, we impose that exporters and non-exporters pay the same returns to observables ($\beta^{NE} = \beta^{E}$). Consider the following decomposition of the log-wage difference between exporters and non-exporter, where $\bar{x}$ denotes within group average of x:

$$\begin{aligned} E(w_{it}|E) - E(w_{it}|NE) =& \bar{w}_{it}^{E} - \bar{w}_{it}^{NE} \\ =& \underbrace{\left[\bar{x}_{it}^{E} - \bar{x}_{it}^{NE}\right]}_{(1)} \beta + \underbrace{\left\{\bar{\psi}_{j}^{E} - \bar{\psi}_{j}^{NE}\right\}}_{(2)} \\ &+ \underbrace{\left\{\bar{\theta}_{i}^{E} - \bar{\theta}_{i}^{NE}\right\}}_{(3)} + \underbrace{\left\{\bar{\tau}_{ij}^{E} - \bar{\tau}_{ij}^{NE}\right\}}_{(4)} \end{aligned} \tag{3}$$

The first term describes differences in observable firm and worker components, as well as differences in general conditions (year dummies). The last three terms describe the differences in unobservable firm,

[15] See also (Card et al., 2018) for further discussions and analysis of this assumption. Compared to the AKM model, our model relaxes the exogenous mobility assumption. This happens in two dimensions. First, workers are allowed to react to the export status of the firm and second, observable characteristics of matches may be correlated with residual match effects (time-invariant for the duration of the match), generating an otherwise "endogenous" decision to move. See also Krishna et al. (2014) who use a similar assumption and discuss exogenous mobility.

worker and residual match components. The decomposition allows us to assess the importance of each component and thus describe which sources are important determinants for the average exporter wage gap.

The results from the decomposition are shown in Table 3. Out of the total exporter wage gap of 3 percent, different observable time-varying characteristics account for 22 percent of the gap. Firm effects account for 94 percent, while worker effects have a negative contribution to the overall gap of -16 percent. In the table we also show to what extent the differences in worker and firm effects can be related to time-invariant observables. Only a small part of the firm fixed effects gap can be explained by industry differences. Decomposing worker fixed effects into time-invariant observables (gender and education) makes the unexplained part of the worker fixed effects gap more negative. The R^2 from the regressions of worker and firm fixed effects on gender-education and industry dummies, respectively are relatively low. It is only 0.04 for the firm fixed effects regression meaning that industry differences cannot account for the observed firm differences in wages. The R^2 for the worker fixed effect regression is 0.27 meaning that the combination of gender and education does account for some of the worker heterogeneity, but far from all of it. Finally, by examining the components of $\left[\bar{x}_{it}^{E} - \bar{x}_{it}^{NE}\right]\beta$ we find that exporters have both higher worker and firm components (0.006 and 0.004, respectively) consistent with the descriptives outlined above.

As a robustness check we examine the importance of the definition of an exporter. In our analysis we have defined an exporter in a certain year as any firm with any export within that year. However, this means that firms that sell even a small amount outside EU borders will be termed "exporter" on the same terms as large firms with large exports. In the appendix (Section B.4) we show that our results are very robust across the different definitions of export status. In particular, we change the minimal requirement for being an exporter from 0 to 100.000 DKK per year and 500.000 DKK per year. The results do not change qualitatively.[16]

Even though our focus is different, it is also interesting to compare our results to the literature more focused on the causal estimate. We do this by calculating the change in the firm effects when a firm changes export status. This would be the equivalent of a local average treatment effect in our setup. In total 6 percent of all firms change export status and the total number of observations in these firms (i.e. worker-years spent in the firms changing status) account for 16 percent of all observations. The total employment-weighted firm fixed effects gap for firms changing export status is 0.010, which is smaller than for the whole sample.[17] Note, that our analysis is designed to combine both the fact that the direct effect of export seems

[16]We have also tried to define exporting firms as the firms, where the exports are e.g. 50 percent or more of the total revenue and non-exporters as those that are not exporting at all. Our results are qualitatively the same across all different specifications. We do not show these in the paper, but they are available upon request.

[17]We stress that the conditions under which the parameter has a causal interpretation (and thus can be interpreted as the average treatment effect of the treated (see Fortin et al. (2011))) are quite strict (firms decision to change export status has to be conditionally

Table 3: Exporter Wage Gap Decomposition

	In Equation (3)	Gap	%
Total		0.031	100%
Time-varying observables	**(1)**	**0.007**	**22%**
- Worker		*0.006*	
- Firm		*0.004*	
- General		*-0.003*	
Firm Effects	**(2)**	**0.030**	**94%**
- Unexplained part		*0.025*	
- Time-invariant observables		*0.005*	
Worker Effects	**(3)**	**-0.005**	**-16%**
- Unexplained part		*-0.008*	
- Time-invariant observables		*0.003*	
Match Effects	**(4)**	**0.000**	**0%**

The table shows the decomposition of the raw export wage premium decomposed into factors accruing to observable characteristics, firm fixed effects, worker fixed effects and match effects. Time-varying worker observables are age squared and age cubic. Time-varying firm observables are firm age squared and firm age cubic. Time-varying general observables are year dummies. See Equation (3) for the decomposition.

very limited, while there still exists a substantial and economically important unconditional exporter wage gap. Also, the relatively small direct effect is in accordance with the literature using exogenous variation to estimate causal effects, see e.g. Hummels et al. (2014) and Macis and Schivardi (2016). We view our results as complementary evidence and a small contribution to the literature focused on causal effects of exporting.

3.1.2 Decomposition with Export-Specific Returns

Above we assumed that returns to observable characteristics are the same in the export and non-export sector ($\beta^{NE} = \beta^{E}$). However, if e.g. some skills are more valuable in the export sector this could lead to higher returns. We therefore estimate an augmented version of Equation (1), where observables have export-specific returns. The model structure now allows for a decomposition of the part of the exporter wage gap accruing to the observables into a part coming from differences in composition and a part coming from differences in returns (Oaxaca, 1973; Blinder, 1973):

$$\begin{aligned} E(w_{it}|E) - E(w_{it}|NE) =& \bar{w}_{it}^{E} - \bar{w}_{it}^{NE} \\ \stackrel{u.e}{=}& \left\{\bar{\psi}_j^E - \bar{\psi}_j^{NE}\right\} + \left\{\bar{\theta}_i^E - \bar{\theta}_i^{NE}\right\} + \left\{\bar{\tau}_{ij}^E - \bar{\tau}_{ij}^{NE}\right\} \\ &+ \left[\bar{x}_{it}^E - \bar{x}_{it}^{NE}\right]\beta^E + \bar{x}_{it}^{NE}\left[\beta^E - \beta^{NE}\right] \\ \stackrel{u.ne}{=}& \left\{\bar{\psi}_j^E - \bar{\psi}_j^{NE}\right\} + \left\{\bar{\theta}_i^E - \bar{\theta}_i^{NE}\right\} + \left\{\bar{\tau}_{ij}^E - \bar{\tau}_{ij}^{NE}\right\} \\ &+ \left[\bar{x}_{it}^E - \bar{x}_{it}^{NE}\right]\beta^{NE} + \bar{x}_{it}^{E}\left[\beta^E - \beta^{NE}\right] \end{aligned} \tag{4}$$

exogenous). In our setting firms are allowed to change export status based on the potential firm fixed effects and its workforce composition, but not on anything else. Nevertheless this source of variation is still an interesting feature of the data we have at hand and gives an idea about the general effect of a change in export status.

where *u.e* and *u.ne* refer to export and non-export returns, respectively. If returns to observables do not differ, the last term in each formulation will simply be zero. We present both representations (i.e. *u.e* and *u.ne*) such that the relative importance of composition and returns can be credibly bounded. In order to separate the effect of differences in fixed effects and differences in returns, we must impose an appropriate normalization to obtain an Oaxaca-Blinder decomposition of the wage gap, which is invariant to a change of reference group (e.g. the omitted year). This is the well known problem associated with the so-called "detailed decomposition" (in our setup this means distinguishing between unobservables and differences in returns) discussed in e.g. Jann (2008) and Fortin et al. (2011). [18]

The results of the decomposition are presented in Table 4. Worker fixed effects still contribute negatively to the overall gap by -18 percent. Differences in observables account for 15 (*u.e*) to 33 (*u.ne*) percent of the gap, while differences in return account for -17 (*u.e*) to -35 (*u.ne*) percent of the gap. Firm effects are still the main driver of the gap accounting for 120 percent. Accounting for industry is not sufficient to remove the gap accounted for by firm fixed effects and still leaves a gap of 97 percent. These results favor Melitz (2003) type models with some form of rent-sharing in the sense that exporting firms simply seem to be better based on unobservables and that this leads to higher wages. The combination of higher firm fixed effects in exporting firms and lower returns to observables implies that exporting firms on average pay higher wages, but compensate workers less based on observable characteristics. For instance, while both workers and firms in the export sector are older on average (see Table 2) the returns to age are lower in this sector.

Overall our results therefore show similar patterns as what was found earlier. Differences in firm pay, unrelated to observables, seem to be the main driver behind the exporter wage gap. But, the high firm fixed effects could arise due to self-selection into export of workers, who are particularly productive on average in this sector. Our current model will interpret this as higher firm effects in exporting firms because worker effects are constant across export status. To accommodate this concern, Section 4 augments the model to allow for export-specific worker fixed effects.

[18] The intuition is that the export wage gap in the reference group will be captured in the job spell fixed effects. Thus if, for example the gap varies over time, what is attributed to year dummies and job spell effects, respectively, will vary conditional on the choice of reference year. This is an undesirable feature and we therefore impose the normalization discussed in Jann (2008) and Fortin et al. (2011) to solve this. In particular we estimate the model with a constant and omitting one year dummy and one job spell fixed effect. We then standardize each set of fixed effects (year fixed effect and job-spell fixed effects under export and non-export, respectively) such that their sample weighted sum is zero. This ensures that we get the same constant regardless of the particular choice of which job spell fixed effect, μ_{ij}, and which year effect we omit. The constants we obtain from the regressions are added back to the job spell fixed effects. For further details on the estimation procedure, see the Appendix, Section A.1. Note that the normalization only affects which part of the gap will be explained by the job spell fixed effects, i.e. μ_{ij}, and therefore the worker, firm, and match effects, and which part will be explained by differences in returns, i.e. $\bar{x}_{it}^{NE}\left[\beta^{E}-\beta^{NE}\right]$ or $\bar{x}_{it}^{E}\left[\beta^{E}-\beta^{NE}\right]$. The "explained" part of the decomposition, i.e. $\left[\bar{x}_{it}^{E}-\bar{x}_{it}^{NE}\right]\beta^{E}$ or $\left[\bar{x}_{it}^{E}-\bar{x}_{it}^{NE}\right]\beta^{NE}$, is the same independent of the normalization (Fortin et al., 2011).

Table 4: Exporter Wage Gap Decomposition

Oaxaca-Blinder	*u.e.*		*u.ne.*	
	Gap	%	Gap	%
Total	0.031	100%	0.031	100%
Explained by observables	0.005	15%	0.011	33%
Unexplained due to returns	-0.005	-17%	-0.011	-35%
Firm Effects	0.038	120%	0.038	120%
Worker Effects	-0.006	-18%	-0.006	-18%
Match Effects	0.00	0%	0.00	0%

The table shows the Oaxaca-Blinder decomposition of the raw export wage premium decomposed into factors accruing to characteristics, firm fixed effects, worker fixed effects, match effects and differences in return (see Equation (4)). Observables include year dummies, worker age squared and cubic, and firm age squared and cubic. Note that as in the decomposition with same returns (Equation (3)) the worker and firm fixed effects terms can be decomposed further into the effect of time-invariant observables and residual fixed effects. The effect of time-invariant observables is similar to the case with same returns (see Table 3). Therefore, to keep things simple we leave this part of the decomposition out. A similar note holds for the decomposition of observables into worker and firm side variables (see Table 3).

4 Export-Specific Worker Skills

In this section we augment the model used in Section 3.1.2 to allow for export-specific returns to unobservable worker characteristics. This is relevant if different time-invariant worker skills and traits are valued differently in the export and the non-export sector. The baseline model featured export-specific firm fixed effects, but not export-specific worker effects. The extended model allows for both export-specific worker and firm fixed effects to separate out the role of self-selection on unobservables of workers into the export sector.

4.1 Normalization of the Firm Effects

We do not make any changes in the first step of the estimation outlined in Equation (1), i.e. the job spell fixed effect model is the same as in Sub-section 3.1.2 where we allow returns to vary with export status However, when decomposing the job-spell fixed effects we allow both firm and worker fixed effects to vary across export status. This corresponds to estimating Equation (2) separately on the export and the non-export samples. However, in doing so, we run into the problem of disconnected sets. Connected sets are central to identification of firm and worker fixed effects, see (Abowd et al., 2002). Intuitively, in this setup both workers and firms are treated as different entities, when the employing firm changes export status. Thus, the set of exporting firms and the set of non-exporting firms can never be connected by the same worker visiting both types of firms as the worker is viewed as two different workers depending on the export status. This implies that there is no immediate way to anchor the fixed effects across these two samples and thus meaningfully compare e.g. firm fixed effects under export and non-export.[19]

[19] The level of the fixed effects in a given sample is always set relative to a normalization (e.g. the last firm effect in the sample is normalized to 0 or the grand mean is taken out). This is not a problem in one connected sample as the normalization applies to both export and non-export firms. But in two different samples we will need two different normalizations. A natural thought is to normalize

The problem of disconnected sets is very similar to the one faced by Card et al. (2016) in the context of different wage schedules for men and women. We follow the solution procedure that they employ, and use the least productive firms in each sample to anchor the fixed effects. In particular, we assume that firms with value-added per worker below a certain estimated threshold create so little surplus in excess of the workers contribution, that they are unable to pay any additional rent to their workers, i.e. we normalize the firm fixed effects below this threshold to be the same across exporting and non-exporting firms. We provide some evidence in favor of this assumption below. One interpretation of this assumption is that in low productive firms the surplus from a match is so low that workers are just paid their marginal product, i.e. return to their observables and export-specific worker fixed effect. However, in more productive firms the match generates a surplus which the worker can capture a part of through bargaining. These additional rents show up in the wage regression as high firm fixed effects and therefore firm fixed effects may differ with export status.

The procedure to anchor the firm fixed effects requires information on value added for firms. In addition, the procedure requires selecting the largest connected set of exporting and non-exporting firms, respectively. We therefore use a subset of the full sample referred to as the value-added (VA) sample. See Table (1) for descriptives of this sample. The VA sample is smaller than the full sample, which we have used until this point. Roughly 95 percent of the lost observations are due to missing or poor quality value-added data. Only 5 percent of the lost observations are due to lack of connectivity. The unconditional exporter wage gap is 0.028 in the VA sample, which is very comparable to the gap of 0.031 in the full sample. Furthermore, the results presented so far are the same for the VA sample which we will show in Table 5.

To determine the set of normalizing firms we use a data-driven approach. In particular, following Card et al. (2016), we run a series of regressions using the firm fixed effects ($\hat{\psi}_j^h$) from the two non-connected sets (separate AKM regressions in the export and non-export samples) and average firm value added ($\bar{va}_j$).[20] We then regress $\hat{\psi}_j^h$ on export status dummies (c^h) and export status specific returns to the value-added fixed effects, $\bar{va}_j$, where firm fixed effects under a varying threshold τ are set to zero (importantly the threshold is independent of export status). That is, for different values of τ and for each $h = \{E, NE\}$ we estimate

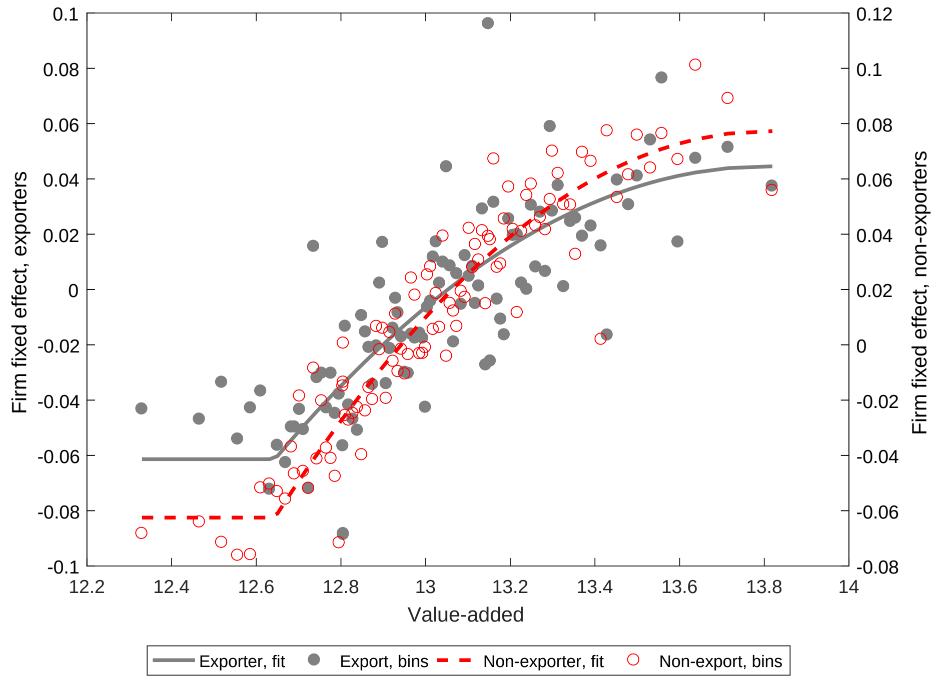
$$\hat{\psi}_j^h = c^h + \gamma^h \max\{\bar{va}_j - \tau, 0\}.$$

c^h is the intercept of the regression for $h = \{E, NE\}$, while γ^h is the slope. τ is the value under which

the same firm in each sample, but this implies that the "normalized" firm has the same firm fixed effect under export and non-export and this normalization is therefore not innocent and may bias the analysis. Below we argue that a set of firms can be used for this normalization such that these firms have the same firm fixed effect under export and non-export.

[20] Note that the value-added we use is detrended in logs such that the mean log value added is the same for all years. For a more detailed step-wise explanation of the normalization procedure see Appendix A.3.1.

Figure 1: The Relation between Value Added and Firm Fixed Effects



The figure shows the mean firm fixed for different percentiles of the value-added distribution for both exporters and non-exporters. The solid lines plot the best square fit for exporters (E) and non-exporters (NE).

value added per worker does not influence the wage the worker receives through the firm fixed effect. We select the τ that minimizes the root mean squared error. Note that the regressions are employment weighted such that an observation is a worker-year. We end up with using a quadratic specification, as it gives the most appropriate normalization (the lowest RMSE, see also Appendix A.3.1). Figure 1 shows the mean un-normalized firm fixed effects for different percentiles of value-added together with the quadratic fit plotted for exporters and non-exporters, respectively. [21]

The figure also illustrates that the economic intuition behind the procedure works very well in our data. Note that if this would not be the case the kink point τ (in the regression outlined above) would simply be set at the smallest firm fixed effect in which case the figure would show that the prediction is not borne out by the data. Therefore, the economic assumption that firms with very low value-added per worker have no surplus to share with their workers, is essentially testable in the data - and not rejected in our case. In the figure this is seen from the fact that for low value added firms there is no relation between value-added and firm fixed effects. This also underlines the importance of our data driven approach (see step 4a described

[21] Note that as firm effects are un-normalized in Figure 1 the level of the firm effects cannot be compared (they can simply reflect that very different firms, not necessarily due to exporting, constitute the normalizing firm (with firm fixed effect set to 0) in the two samples), see also Card et al. (2016). The key point of the figure is that a set of firms with low value added have very similar fixed effects (fixed effects do not change with value added for these firms), this linked with the assumption that such low value added firms do not have any surplus to share neither as exporters nor non-exporter implies that we can anchor fixed effects across samples (strictly speaking we only need that these firms do not change their firm effects with a change in export status, a firm effect different from 0 is allowed).

Table 5: Export-Specific Worker Skills

	Baseline (1)		VA-norm. (2)		Transport norm. (3)	
Raw gap	0.028	100 %	0.028	100 %	0.028	100 %
Observables	0.007	25 %	0.007	25 %	0.007	25 %
Firm effects	0.027	96 %	-0.010	-39 %	0.003	11 %
Match effects	0.000	0 %	0.000	0 %	0.000	0 %
Worker Effects	-0.006	-21 %	0.031	111 %	0.018	64 %

The table shows the decomposition of the wage gap into worker, firm, match and observables for different models. The baseline model (Column (1)) refers to the model from Section 3.1.2. The VA-norm. (Column (2)) refers to the model using the value-added normalization, and the Transport norm. (Column (3)) refers to the model using the transport sector normalization (both normalizations are introduced in Section 4). All three models, Column (1)-(3) are based on the VA sample (see Table 1).

in Appendix A.3.1). The normalization applies to 8.2 percent of the non-exporting firms and 2.9 percent of the exporting firms. It accounts for 4.5 percent of observations in non-exporting firms and 0.9 percent of observations in exporting firms.[22]

As a robustness test we also apply a normalization where we normalize firms in the transport sector to have firm fixed effects of zero. The reason for selecting the transport sector is that the unconditional exporter gap is almost zero in this sector. A reason for the absence of a gap in this sector could be because firms within this sector cannot afford to pay an export wage premium. The basic economic intuition is the same as in the argument by Card et al. (2016): There exists a set of firms for which rent extraction by workers is impossible, because there is no surplus to extract. The results for the different normalizations are seen in Table 5 and discussed in the following section.

4.2 Results

In Table 5 we show the results from the two normalizations of the firm effects and compare them to the previous results, where we did not allow the worker fixed effects to change with export status.

Column (1) in Table 5 shows the decomposition of the model from Section 3.1.2 into observables, firm fixed effects, match fixed effects and worker fixed effects for the VA sample. The results in Column (1) are very similar to those estimated on the Full sample (see Table 3) confirming that moving to the VA sample does not alter our main conclusions thus far. Columns (2) and (3) show the results from the model where workers are allowed to have export-specific unobservable skills under two different normalizations: the value-added and the transport sector normalizations. The qualitative conclusion is similar in both Column (2) and (3): What the baseline model interprets as higher firm fixed effects is instead the result of workers in export sector having higher worker fixed effects on average in this sector compared to the non-export sector. We will explicitly show this in Table 6.

[22]Note that the exporting firms lie systematically above the flat line in Figure 1. This is because of the two major outliers, both of which come from very large firms. The graph suggests that the value added per worker measure we use for these firms is most likely too low.

If we compare Column (1) to Column (2), worker effects account for -0.006 and 0.031 for the baseline and value-added normalization, while firm effects account for 0.027 and -0.010. Thus, worker and firm effects play opposite roles in the two models. In the baseline model, where worker fixed effects cannot change with export status, firm effects are the main driver of the gap while worker effects close the gap. This implies that exporting firms are high wage firms, but employ low wage workers. This conclusion is, however, reversed in the model with export-specific worker skills. In that model, workers in the export sector are paid more because of their unobserved skills, while exporting firms actually pay a lower wage overall. If we interpret wages as a proxy for productivity this implies that it is not exporting firms that are positively selected. Instead, workers in exporting firms have skills that are more valuable in the export than in the non-export sector. For the counterfactual experiment of moving a random worker from a random non-exporter to a random exporter, this is extremely important. If the firm fixed effects are generally higher in the export sector, we would expect the worker to get a higher wage. However, if the exporter gap is driven by differences in export-specific worker fixed effects, we cannot make such a claim: The workers are probably already selected into the sectors, where they have a comparative advantage. The results from the normalization to the transport sector shown in Column (3) deliver the same results qualitatively.[23] The intuition behind the change in the result is fairly simple. Recall, that the firm effect is identified from workers moving between two firms. In the estimation where we did not allow for export-specific worker effects the higher wage that worker's on average received transiting from a non-export firm into a specific exporting firm was automatically interpreted by the model as higher firm effects in the exporting firm. However, with export-specific worker effects this no longer happens automatically; instead the change in wages may now reflect export-specific worker effects or export specific firm effects. Distinguishing between the two mechanisms is possible due to the scope of the variables: workers always get paid their worker effects regardless of the firm whereas the firm always "pays" its firm effect to all its workers. Hence, to distinguish between the two sides of heterogeneity the distribution of wage changes in the firm (and the normalization above) becomes important:; if all workers who make the same transitions from a non-exporting into an exporting firm also experience similar wage changes then the model interprets this as evidence in favor of firm effects. Oppositely, when there is heterogeneity in the wage changes for workers moving from the non-exporting into the exporting firms the model interprets this as evidence in favor of exporter-specific worker effects which the worker then always gets in exporting firms (below we provide some evidence of this heterogeneity).[24]

[23] The same results hold for the linear or cubic fits of the value-added regressions which we do not report here.

[24] Interestingly, Macis and Schivardi (2016) find results which are similar in their analysis of the changes in the wage structure in response to an Italian currency devaluation. Their analysis shows that in a version with fixed worker effects the AKM regression attributes a too large part of the change in wages after the devaluation to a change in firm fixed effects (a rent sharing effect in their framework). However, when worker effects are allowed to change this rent sharing channel is much less important. Note that our

For further illustration we use the sub-sample of workers, who work in both the exporting and non-exporting sector during our sample window and investigate the change in the worker effect associated with a change of export status. For this sub-sample of workers, which amounts to 27 percent of all worker-year observations (or 22 percent of all workers) in the VA sample, we can decompose the change in worker effects as follows:

$$
\begin{aligned}
\left\{\bar{\theta}_i^E - \bar{\theta}_i^{NE}\right\} =& E\left[\theta_i^E | E\right] - E\left[\theta_i^{NE} | NE\right] \\
\overset{u.e}{=}& E\left[\theta_i^E - \theta_i^{NE} | NE\right] + \left\{E\left[\theta_i^E | E\right] - E\left[\theta_i^E | NE\right]\right\} \\
\overset{u.ne}{=}& E\left[\theta_i^E - \theta_i^{NE} | E\right] + \left\{E\left[\theta_i^{NE} | E\right] - E\left[\theta_i^{NE} | NE\right]\right\}
\end{aligned}
\tag{5}
$$

where *u.e* and *u.ne* refer to under export and non-export returns, respectively. The conditioning ($|E$ or $|NE$) simply weights the worker effects by the number of worker-year observations that the worker is observed in that sector. Consider the decomposition under export: $E\left[\theta_i^E - \theta_i^{NE} | NE\right]$, it gives differences *in returns* to unobservables between exporting and non-exporting firms weighted by the number of worker-years in the non-export sector. Thus, it answers the question: How much higher wages would a worker in non-export on average get if he moved to the export sector due to comparative advantage in this sector. $\left\{E\left[\theta_i^E | E\right] - E\left[\theta_i^E | NE\right]\right\}$ shows *composition* differences, i.e. whether high wage export workers spend more time in the export or non-export sector. The decomposition of the gap under non-export returns has a parallel interpretation.

In Table 6 we first repeat the decomposition in Table 5 for the sub-sample of workers that at some point worked in both the export and non-export sector. The main conclusions are the same. We then focus on the new decomposition of the worker effects shown in Equation (5). Again, we focus on Column (2), but the results hold across the different normalizations. In Column (2) the total worker fixed effect gap is 0.031 for workers who change export status. Interestingly, the gap is driven entirely by differences in worker fixed effects, i.e. the first term of the decomposition in Equation (5) under both the export and non-export distributions. This term accounts for a gap of 0.034-0.040, while there is a small adverse composition effect, as in the model without export-specific returns to unobservables. Thus, the model with export-specific worker fixed effects may have similar implications for the workers who do change sector, as the model without export-specific effects: They get higher wages on average in exporting firms. But importantly, whereas the baseline model extrapolates this to hold for all workers by capturing it in the firm effects, the model with export-specific worker effects attributes this wage increase to differences in worker effects for

setup is different as we allow the market value of worker unobservables to differ depending on the export status of the firm and not just as a response to a change in the economic environment (currency devaluation). We thereby provide further direct evidence on the importance of export-related skills. Note that a meaningful quantification of export-related skills or export-specific firm and worker effects in either setup (i.e. in the devaluation setup or in ours) crucially relies on the assumption about the normalizing firm.

Table 6: Export-Specific Worker Skills for movers

	Baseline (1)	**VA-norm. (2)**	**Transport norm. (3)**
Raw gap	0.012	0.012	0.012
Observables	0.007	0.007	0.007
Firm effects	0.011	-0.026	-0.014
Match effects	0.000	0.000	0.000
Worker Effects	-0.006	0.031	0.019
Worker effects decomposition			
Return under u.e		*0.034*	*0.021*
Composition under *u.e*		*-0.003*	*-0.003*
Return under *u.ne*		*0.040*	*0.027*
Composition under *u.ne*		*-0.008*	*-0.008*

The table shows the decomposition of the wage gap into worker, firm, match and observables for different models. The table focuses on the decomposition of the worker component. The table is based on the sub-set of workers in our VA sample (see Table 1), who during the sample window work in both the export and non-export sector. The model in (1) is presented in Section 3.1.2. The model in (2) and (3) is presented in the Appendix C.2.

those working in the export sector.

We interpret the difference between the worker effect in export and non-export as a sign of a comparative advantage leading to differential returns in one sector over the other. We do not find evidence suggesting that the comparative advantage gap is linked to any specific industry or occupation. Figure 3 in the Appendix shows the mean comparative advantage for different industries and occupations. The health industry is the only industry with a very large gap, but this industry has few observations. We therefore conclude that the distinction between the exporting and non-exporting sector is not just capturing industry or occupational differences, but something more fundamental. The same picture emerges when we compare the 20 % worse workers in terms of comparative advantage (in this case disadvantage) to the 20 % best. The two groups are very similar in terms of observables. The best workers are slightly older, more likely to be males and work in top management or manual labor, but none of the differences are large and it shows that the usual sub-group investigations of the impacts of export will miss substantial heterogeneity in the returns to export.[25]. The largest difference is that the best workers have larger worker effects under export and smaller worker effects under non-export - this highlights that the "best" workers really have comparative advantage and are not just better workers.

The estimation setup in the extended model can be viewed as fairly complex and resting on multiple assumptions including the assumptions behind the normalization of firm effects. As a sanity check of the economic content of our finding we show that workers react to the differential returns by selecting into the sector, where they are relatively best. We provide two pieces of evidence below: First, Figure 2 shows the sample likelihood of working in the export sector conditional on the size of the comparative advantage in the export sector ($\theta_i^E - \theta_i^{NE}$). The figure shows clear evidence of workers exploiting their comparative ad-

[25]Macis and Schivardi (2016) also document substantial heterogeneity in the returns to export which is not captured by common observables.

Table 7: The Effect of a Comparative Advantage on Transitions

Dependent variable:	Export Status of Destination Firm
Standardized Values of Gap	0.0109***
	(11.75)
Export status of Origin Firm	-0.314***
	(-162.95)
Destination Firm FE	-0.0580***
	(-59.53)
Origin Firm FE	0.00230*
	(2.37)
Constant	0.614***
	(465.56)
Observations	241,868

Note: The table shows the effect of the worker fixed effect gap on the export status of the destination firm conditional on a job-to-job transition with a number of control variables. The Export Status variable takes the value 1 if the firm is an exporter. Origin and Destination Firm FE are standardized. t statistics in parentheses, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

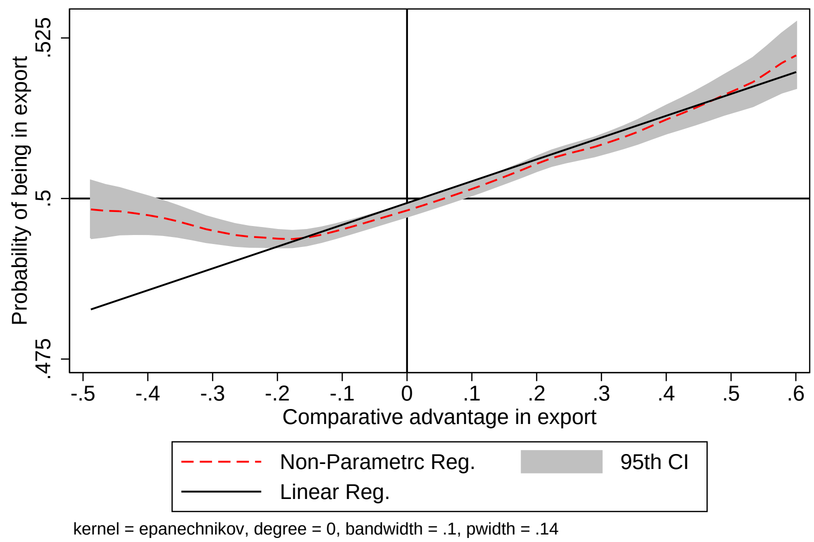
vantage. Workers with a positive comparative advantage in the export sector also tend to work significantly more in the export sector.[26]

Second, instead of looking at the stock of employed workers, we look at the job-to-job transitions in our sample and analyze whether workers who have a large comparative advantage in a sector are more likely to move to this sector. I.e. we investigate whether the transitions we actually observe are directed towards the sector, where the worker has a comparative advantage. We regress the export status of the destination job on the size of the comparative advantage, where the gap has been standardized to have mean zero and standard deviation one. We control for origin firm fixed effect and destination firm fixed effect, since these factors are allowed to affect mobility in our model. In addition, we control for export status of the origin firm. This is important, as we are conditioning on a job move as well as workers having worked in both export and non-export firms. This combination will make cross-sector moves over-represented in the sample. Table 7 shows the result of this regression. A comparative advantage in the export sector statistically significantly increases the probability that the destination firm is an exporter. In particular, a one standard deviation increase in comparative advantage increases the probability that the receiving firm is an exporter with 1.1 percentage point. Furthermore, the effect is larger than the effect of a corresponding one standard deviation increase in the origin firm's firm fixed effect and 20 % of the effect of the receiving firm's firm fixed effect. Thus, it has real economic content.

In this section we showed that workers have unobserved skills that have sector-specific (differential) returns. The worker-specific differential returns is the most important driver of the overall exporter wage gap. Workers react to differential returns by selecting into the sector, where they are relatively best.

[26]Note that the fact that our sample window covers only 6 years and that a worker needs to be present in both export and non-exporting firms to be included in this graph makes the scope for large differences in likelihood limited. Therefore it is remarkable that we get such a clear and significant difference.

Figure 2: Probability of Working in the Export Sector Conditional on Comparative Advantages



The figure shows the probability of working in the export sector as a function of the size of the comparative advantage ($\theta_i^E - \theta_i^{NE}$). The table is based on the sub-set of workers in our VA sample (see Table 1), who during the sample window work in both the export and non-export sector. We trim the ($\theta_i^E - \theta_i^{NE}$)-gap at the 1th and 99th.

5 Non-Additive Worker and Firm Effects

So far we have shown that allowing for flexibility in the interaction between layers of heterogeneity, in particular worker effects, firm effects, and export status, is important for a credible prediction about the drivers behind the exporter wage gap. To this point a key assumption in the analysis has been the additive link between worker, firm, and match effects. Also, the assumption that match effects should be orthogonal to worker and firm effects might explain why they so far have not been important in explaining the exporter wage gap. Finally, the AKM model has some known problems such as limited mobility bias and the fact that the fixed effects are not consistently estimated except if workers are observed for an infinite amount of time (see e.g. Andrews et al. (2012)). In this section we investigate to which extent our results are driven by these assumptions and the particular model that we have used. We adapt a novel framework recently developed by Bonhomme et al. (2017) (henceforth BLM) to our setting. In the BLM framework wage profiles can be non-linear and the result of unrestricted interactions between particular worker types and firm classes. Note that in the AKM framework studied until now such interactions are not allowed: the wage is simply the addition of worker and firm components and there is no additional term dependent on the worker-firm match except the so-called match effect, which was assumed orthogonal to the worker and firm effects. In the BLM framework workers are allowed to act on the match specific effects and this may lead to specific composition patterns. In our implementation workers can have different transition patterns depending on the firm's export status. Thereby we allow exporting firms of a certain class to be able to attract particular worker types with a different probability than their non-exporting counterparts. This may happen either because these workers are simply better workers or because there is some complementarity between firm and worker types. We also allow exporters and non-exporters to pay different wages to a specific worker type instead of paying the same to all workers as is the case with the firm fixed effect in the AKM model.

The setup of the BLM model is as follows. We assume that the world is well described by a set of firm classes K and worker types L (later we analyze whether our results are robust to the choice of K and L). The fact that there is a fixed number of firm classes and worker types makes the estimated parameters consistent. All interaction effects result from particular wage schedules for each worker type-firm class pair. Differences in wage schedules can arise for multiple non-modeled reasons such as complementarities in production, preferences, bargaining power, etc. The BLM procedure assumes that firm class membership can be determined in an initial step prior to estimation and therefore it is treated as observed in the final estimation. Worker types, on the other hand, are treated as unobserved data, and group membership is estimated.[27] As in the AKM model wage changes associated with job moves are crucial for identification in the BLM model. Differences in wage schedules are identified by workers of different types moving to and from each firm class. This is similar to the notion of connected sets in the AKM framework. If worker types are "different" in different firms (i.e. worker and firm effects are non-additive), we must require that firms are connected for each worker type. This would be impossible if not for the assignment of workers and firms into a number of groups. Intuitively, this allows us to observe the "same" worker (worker type) making multiple transitions at the same point in time, hence creating additional degrees of freedom needed to estimate interaction effects between worker and firm types. The grouping also implies that we only need two years of data to estimate the BLM model, which also relaxes the implicit assumption that firm and workers fixed effects do not change in our sample window.

We assume that log-wages are drawn from a firm-class/export-status/worker-type specific (i.e. (k, h, l)-specific) distribution.[28] Thus,

$$w_{ijt} = \lambda_{k(j),h(j),l(i)} + \sigma_{k(j),h(j),l(i)}\delta_{ijt}$$

where w_{ijt} is the log wage of worker i in firm j at time t. Sometimes we will omit subscript j as the structure

[27]This essentially amounts to treating workers' classification as a random effect (i.e. we get a probability distribution across types), whereas firm are treated as fixed effects (we can "fully" recover the type). BLM argue that this asymmetry can be justified by the observation that in the data we see many "observations" for each firm, whereas the number of spells for each worker is limited.

[28]We do not include observables in the model as this separation is not essential for our main purpose, namely to assess whether the exporter wage gap comes from different workers, different firms or different matches. In fact the BLM model does not include any direct effect of observable characteristics on wages. Hence, in the most direct formulation it is not possible to make the distinction between observed and unobserved differences. Differences in worker type could therefore also partly capture cross sectional differences in experience (or age), and firm class could be correlated with firm age. Note, that the BLM estimator only uses two periods of data, so time-trends and experience accumulation are of limited importance in the sample.

of our data is already completely specified by the i, t-notation. $h(j)$ is the mapping from firm identifier to the firm export status, which is either $h(j) = \{E, NE\}$. $k(j)$ is the mapping from firm identifier to firm class k. Similarly, $l(i)$ is the mapping from worker identifier i to worker type l. The mapping functions are important as we work with a limited number of worker and firm types. Hence, several firms in the data map into the same firm type and likewise for workers. We will sometimes use the notation $j(i, t)$ to map worker i to her employer j at time t. $\lambda_{k,h,l}$ and $\sigma^2_{k,h,l}$ denotes the firm-class/export-status/worker-type specific means and variances of wages. Following BLM, we assume that δ_{ijt} follows a standard normal distribution. These assumptions imply serial independence in w_{ijt} once conditioning on worker type, firm class and export status. Let $\{O, D\}$ denote origin and destination firm class for movers. We assume that, conditional on worker type, origin firm class k^O_t and origin export status h^O_t, the decision to move and where to move to (for those that move), i.e. $\{k^D_{t+1}, h^D_{t+1}\}$, is independent of the current draw of w_{ijt}. Thus, we allow transitions and wages to explicitly depend on the export status of the firm.

Let $m_i = 1\{j(i, t) \neq j(i, t + 1)\}$ indicate whether a worker moves or not and let $J_{it} = 1$ if the worker was employed at time t and zero otherwise. Note, we include moves to and from non-employment. Using the above assumption, we derive the following log-likelihood function for our two period sample:[29]

$$\begin{aligned}\mathcal{L}(\theta) = &\sum_{i=1}^{N}\sum_{k=1}^{K}\sum_{k'=1}^{K}\sum_{s=0}^{1}\sum_{s'=0}^{1} 1\{k^O_{i1} = k\}1\{k^D_{i2} = k'\}1\{h^O_{i1} = s\}1\{h^D_{i2} = s'\} \\ &\cdot \log\left[\sum_{l=1}^{L} p_{k,k',s,s'}(l)^{m_i}\bar{p}_{k,s,s'}(l)^{1-m_i}\phi\left(\frac{w_{i1} - \lambda_{ksl}}{\sigma_{ksl}}\right)^{J_{i1}}\phi\left(\frac{w_{i2} - \lambda_{k's'l}}{\sigma_{k's'l}}\right)^{J_{i2}}\right]\end{aligned}$$

where h_{it} indicates the export status of the employer of worker i at time t and ϕ is the standard normal density. $p_{k,k',s,s'}(l)$ is the worker type-specific probability of transitioning from firm class k to firm class k' with export status s and s', respectively. $\bar{p}_{k,s,s'}(l)$ is the counterpart for stayers in firm class k. Note that $p_{k,k',s,s'}(l)$ includes the transition probability in and out of unemployment for each firm class. As proposed by BLM we maximize the log-likelihood function using the EM-algorithm (see Dempster et al. (1977)).

We include firms and workers from the VA sample used for estimation of the model in Section 4. Note that the BLM framework only requires two years for estimation. We use the years 2009 and 2010 and set the number of firm classes to $K = 5$ (and an additional unemployed state) and the number of worker types to $L = 5$. We have tested the robustness of our results to these choices. In particular, we have tried using years 2006-2007 (see Table 13 in the Appendix) and we have also estimated the model with $K = L = 2$ and $K = L = 10$ (See Table 12 in the Online Appendix) and the results are very stable. For the initial assignment

[29]Note that Bonhomme et al. (2017) model the joint distribution of wages, w_{ijt}, for job movers over two periods, i.e. the joint distribution of (w_{i1}, w_{i2}). We also include stayers in the estimation as suggested by Lentz et al. (2016) to utilize the full set of wage data.

of firms into firm classes we use the value-added data and classify firms into quintiles of the distribution of average value added per worker. After the initial assignment we interact class membership with export status. Thus, the classification implies that we allow firms with the same productivity to have different wage and hiring policies depending on their export status.[30] We include an unemployment state in the model, keeping workers from the connected set, if they are employed at least one of the sample years. In the case of unemployment in the first period, we require that workers have been in the connected set during the previous two years. In the case of unemployment in the second period, we require the worker to return to employment within the following two years. These restrictions are imposed to avoid issues of long term unemployment and attrition from the labor market.

5.1 Decomposition

If the true wage equation is non-additive as outlined above, AKM would lead to biased estimates which would affect the decompositions above essentially because we risk assigning heterogeneity to the wrong side of the market. Instead the BLM estimation procedure allows us to consistently estimate worker and firm types when the true wage equation in non-additive. The resulting expected wages (λ_{khl}) are match specific. However, we are still interested in the decomposition of the exporter wage gap into worker, firm and match effects, where match effects are now seen as interaction effects as opposed to the orthogonal match effects in the AKM framework. Therefore, as a next step in the analysis, we project the estimated worker, firm, and match type indicators onto λ_{khl}. This allows us to analyze to what extent the non-additive formulation changes the conclusions outlined above.

Before we decompose the wage gap, note that the expected wage conditional on export status $h = \{E, NE\}$ is

$$E\left[w_{it}|h\right] = \sum_{k=1}^{K} P\left[k_{it} = k|h\right] \sum_{l=1}^{L} P\left[l_i = l|k_{it} = k, h\right] \lambda_{khl}$$

We may non-parametrically represent λ_{khl} as

$$\lambda_{khl} = \psi_{kh} + \theta_{lh} + \tau_{klh} \tag{6}$$

where ψ_{kh} represents the firm fixed effect of the class k with export status h, θ_{lh} represents the worker fixed effect of a type l worker in a export status h firm, while τ_{klh} represents the non-linear interaction effect

[30] BLM classify firms based on their within-firm empirical wage distributions (see Bonhomme et al. (2017)). For our agenda using wage information to classify firms while distinguishing them by export status does not seem credible - this essentially amounts to saying that firms are of the same type if they pay the same wages which then assumes that the wage distribution of the firm does not change with export status. As the firm classification is done prior to estimation in the BLM framework we can easily change this and instead group firms by value added per worker. The assumption then is that the wages in e.g. low value added non-exporting firms are a good counterfactual for exporting firms (after having taken into account any selection effects coming form the fact that matches in the firms could be very different).

for a type l worker in a class k firm with export status h. As in the two-way fixed effect model of AKM, we need some normalizations to recover the full set of fixed effects. In the AKM framework we needed one normalization to separate the level of the firm and worker effects - in the current context we need further restrictions. λ_{klh} contains $2 \cdot K \cdot L$ cells or different wage observations (exporter status interacted with worker and firm types), but (6) above involves $2 \cdot K + 2 \cdot L + 2 \cdot K \cdot L$ parameters. Thus, for estimation we need $2 \cdot (K + L)$ normalizations. Below we present two alternative sets of normalizations to recover the worker, firm and match effects, which are designed to mimic the type of decompositions we made in the additive model:

N1: Standard Normalization: In this specification we require the interaction effects to lie around the corresponding worker and firm effects, such that match effects are zero in expectation within a worker/firm type. Thereby, the interaction effects get the appealing interpretation of deviations from additive wages. Furthermore, we assume that workers do not have export-specific returns. These restrictions are therefore similar to those used in the models in Section 3 and 3.1.2. Formally, we normalize the interaction effects τ_{klh} such that for each $k = \{1, .., K\}$ we have $E[\tau_{klh}|k, h] = 0$ within each sector $h = \{E, NE\}$, while for each $l = \{1, .., L\}$ we have $E[\tau_{klh}|l] = 0$. In addition, we assume that workers do not have export-specific returns, i.e. $\theta_{lE} = \theta_{lNE}$. The assumption gives the $2 \cdot (K + L)$ restrictions needed. Lastly, in order to separate worker and firm effects we normalize an arbitrary firm effect to zero. This does not affect the decomposition and amounts to the usual AKM assumption. We refer to Section C.1 in the appendix for a derivation of the full decompositions and wage function. It is worth noting that the restrictions on the interaction effects are weaker than the AKM requirement of orthogonal match effects, which would correspond to assuming $E[\tau_{klh}|k, l, h] = 0$ for each set of k,l and h.

N2: CCK-Style Normalization: In this specification we treat worker, firm, and match effects as if they came from two separate BLM models. In this spirit, we therefore assume that: $E[\tau_{khl}|k, h] = 0$ for each k and h, and $E[\tau_{khl}|l, h] = 0$ for each l and h. This is equivalent to $2 \cdot (K + L)$ restrictions. The restrictions imply that we interpret match effects as deviations from the export-specific combination of worker and firm effects. To separate worker and firm effects we need one normalization within each sector. Here, we will assume that the lowest firm class has so low productivity that their firm effect is zero both when exporting and not. This mirrors the normalizing assumption of Section 4 pioneered by Card et al. (2016) (CCK). As in Section 4 this is an economic assumption that influences the decomposition. The normalization allows us to do a full decomposition. We refer to Section C.2 in the Appendix for the technical details given these normalizations.

Note that both types of normalizations by construction will be limiting the role of match effects in the mean wage decomposition as the normalizations in various ways set the sample weighted mean to zero

Table 8: Decomposition of the Exporter Wage Gap with Non-Additive Wages under Exporter Returns

2009-2010	**BLM N1 (1)**	**AKM (2)**	**BLM N2 (3)**	**AKM (4)**
		under exporter returns		
Total	0.046	0.046	0.046	0.046
Worker	0.015	0.003	0.044	0.040
- Composition			*0.015*	*0.004*
- Return			*0.028*	*0.035*
Firm	0.031	0.028	0.002	-0.009
- Composition	*0.020*	*0.003*	*0.020*	*-0.005*
- Return	*0.011*	*-0.004*	*-0.017*	*-0.031*
Match	0.000	0.000	0.000	0.000
- Composition	*0.000*	*0.001*	*0.000*	*0.000*
- Return	*0.000*	*-0.001*	*0.000*	*-0.001*

The table shows the decomposition of the exporter wage gap from the BLM model estimated on the years 2009-2010 under the assumptions of Section 5.1. Column (1) is based on the restrictions N1 (see Section 5.1), while Column (2) shows the AKM equivalent (the decomposition in Section 3 on the BLM sample). Column (3) is based on N2, while Column (4) shows the AKM equivalent (i.e. the decomposition from Section 4). All decompositions are under the non-exporter returns (the similar decomposition under non-exporter return delivers very similar results see 11 in the Appendix). For the AKM decompositions (Column 2 and 4) some elements of the decomposition are omitted and the detailed decomposition of worker, firm and match effects is only based on the appropriate identifying sub-samples, therefore rows do not sum the total.

within e.g. firm types. This ensures that the decompositions are more comparable to the decompositions in the linear model, but note that workers in the BLM framework are allowed to act on these interaction effects and the estimation allows us to take this composition effect into account.

5.2 Results

Table 8 shows the results from the decompositions. The BLM model explains the unconditional wage gap between the export and non-export sector of 4.6 percent based on the years 2009 and 2010. We present the decompositions under the return of exporters. However, the decompositions under the returns of non-exporters are very similar, see Table 11 in the Online Appendix.

Column (1) in Table 8 presents the results under the Standard Normalization (N1 above), Column (3) presents the results under the CCK-type Normalization (N2 above). Column (2) and Column (4) present the AKM versions of two normalizations, respectively (i.e. the models from Sections 3 and 4, respectively, rerun on the same sample as the BLM model). Looking at Column (1), the gap is explained by the firm component and the workers component. This is qualitatively similar to the AKM results in Column (2). Turning to Column (3) of Table 8 we also get results very similar to the equivalent AKM model in Section 4 shown in Column (4).[31]

The results imply that both worker and firm effects contribute to the total wage gap, while the match

[31] Keep in mind here that a firm type (or worker type) in the BLM framework is not defined at the level of each individual firm as in the AKM model. Hence the decomposition also shows that although one "aggregates" firms to 10 different types (5 non-export and export firm classes respectively) in the data our model is still sufficiently flexible to generate differences between workers and firms within export status and thus capture the important variation in the data.

effect does not matter much. This is of course partly a result of the way we have set up the normalizations above, but Bonhomme et al. (2017) do not find that match effects are important either. Interestingly, the worker component is less important when returns are not allowed to vary with export status seen in Column (1), but allowing for worker's return to export-specific skills increases the component from 0.015 to 0.044. Exporting firms are still positively selected, but in accordance with the results in Section 4, we find that firms actually become less important in export (the firm effects part falls from 0.031 to 0.002). This means that an arbitrary worker would not benefit from working in export everything else equal, only workers with the "right" unobservables will benefit.

Both normalizations imply that some worker type-firm class combinations are more attractive in export than others. This is evidence in favor of our main conclusion: an important part of the exporter wage gap is explained by specific worker types getting higher returns under export than they would under non-export. Again, we stress that this is not a global statement as it does not hold for all workers. Thus, using a very different framework than the more standard additive AKM model, we find qualitatively the same results: Failing to account for differential returns to worker-specific skills in the export and non-export sectors implies that the primary driver of the exporter wage gap is firm selection. However, allowing for the returns to worker-specific skills to differ in the two sectors changes the conclusion dramatically. The primary driver is that workers with skills valuable in the export sector sort into this sector. Our findings can be understood as comparative advantage - certain workers have unobserved skills (for instance language or communication skills) that are either more valuable or used more often in exporting firms, and therefore these workers earn higher wages in such firms. Above we have shown that workers also make transitions based on this.

6 Conclusion

It has been widely documented empirically that workers in exporting firms earn higher wages than workers employed in non-exporting firms. Similarly, there is a large theoretical literature on the differences between exporting and non-exporting firms and their workers. The importance of different sources of heterogeneity in generating the exporter wage gap is likely to not only influence the credibility of different theoretical trade models, but also public opinion on trade liberalization. We apply two different econometric frameworks to asses the different sources generating the exporter wage gap. In the baseline setting we use an AKM-type framework. Controlling for observables as well as worker and firm effects, we find that differences in observable characteristics account for 22 percent of the exporter wage gap, firm fixed effects account for 94 percent of the gap, while worker fixed effects account for -16 percent of the gap. Allowing firms to pay different returns to observed characteristics does not qualitatively alter this conclusion. Actually, the firm effects become even more important in explaining the exporter wage gap.

The conclusion that firm effects are the most important driver of the exporter wage gap is, however, drastically altered when we allow workers to have export-specific returns to unobservables (the worker fixed effect). This addition to the model is important as the baseline model allows firm fixed effect to change with export while workers remain the same. Returns to unobservable skills for workers would differ, if for instance firms valued time-invariant worker characteristics differently in the export and non-export sector. Our results imply that they do. Using the extended model we find that worker effects are the most important driver of the total exporter wage gap, while firm effects have a smaller negative contribution. What looked like higher firm fixed effects in the export sector in the baseline model is actually the result of workers in the export sector having skills that are more valued in exporting firms than in non-exporting firms. We also show that workers with a comparative advantage in the export sector are more likely to be in this sector in the cross-section and are more likely to move to the export sector conditional on making a job-to-job transition. Finally, since the AKM model has various disadvantages, we show that using a non-linear random effects model as suggested by Bonhomme et al. (2017) we obtain qualitatively similar results as in the AKM-style framework.

Our findings of comparative advantage for some workers in the exporting sector is an important conclusion, and it changes the understanding of the exporter wage gap. If the gap results from common firm effects, all workers should receive higher wages when working in the export sector. However, if differences come from exporting firms valuing specific skills or traits among workers more than the non-exporters, only some workers will gain from moving to the export sector. Our findings are strengthened by the fact that workers actually make labor market transitions based on their comparative advantages. It is also an important point of the paper that standard observables explain very little in terms of who benefits from working in the exporting sector. This implies that the heterogeneous returns to export that we have documented above is driven by something inherently unobservable and more fundamental than e.g. education levels or the type of occupation/industry. We leave to future work to investigate this further.

7 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., Kramarz, F., and Margolis, D. N. (1999). High Wage Workers and High Wage Firms. *Econometrica*, 67(2):251–334.

Abowd, J. M., McKinney, K., and Schmutte, I. M. (2015). Modeling Endogenous Mobility in Wage Determiniation. Working Papers 15-18, Center for Economic Studies, U.S. Census Bureau.

Andrews, M., Gill, L., Schank, T., and Upward, R. (2012). High wage workers match with high wage firms: Clear evidence of the effects of limited mobility bias. *Economics Letters*, 117(3):824–827.

Bernard, A. B. and Jensen, J. B. (1997). Exceptional exporter performance: Cause, effect, or both? NBER Working Papers 6272, National Bureau of Economic Research, Inc.

Bernard, A. B., Jensen, J. B., and Lawrence, R. Z. (1995). Exporters, Jobs, and Wages in U.S. Manufacturing: 1976-1987. *Brookings Papers on Economic Activity*, 26(1995 Micr):67–119.

Blinder, A. S. (1973). Wage discrimination: Reduced form and structural estimates. *The Journal of Human Resources*, 8(4):436–455.

Bonhomme, S., Lamadon, T., and Manresa, E. (2017). A Distributional Framework for Matched Employer Employee Data. Working Paper (Revise and resubmit in Econometrica).

Card, D., Cardoso, A. R., Heining, J., and Kline, P. (2018). Firms and Labor Market Inequality: Evidence and Some Theory. *Journal of Labor Economics*, 36(S1):13–70.

Card, D., Cardoso, A. R., and Kline, P. (2016). Bargaining, Sorting, and the Gender Wage Gap: Quantifying the Impact of Firms on the Relative Pay of Women. *The Quarterly Journal of Economics*, 131(2):633–686.

Correia, S. (2016). Linear Models with High-Dimensional Fixed Effects: An Efficient and Feasible Estimator. Working Paper. http://scorreia.com/research/hdfe.pdf.

Davis, D. R. and Harrigan, J. (2011). Good jobs, bad jobs, and trade liberalization. *Journal of International Economics*, 84(1):26–36.

Dempster, A. P., Laird, N. M., and Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of The Royal Statistical Society, Series B*, 39(1):1–38.

DST (2008). Udenrigshandel med varer og tjenester - Kilder og metoder. Statistics Denmark.

Eriksson, T., Smeets, V., and Warzynski, F. (2009). Small open Economy Firms in International Trade: Evidence from Danish Transactions-Level Data. *Nationaløkonomisk tidsskrift*, 147(2):175–194.

Fortin, N., Lemieux, T., and Firpo, S. (2011). Chapter 1 - decomposition methods in economics. volume 4 of *Handbook of Labor Economics*, pages 1 – 102. Elsevier.

Frías, J. A., Kaplan, D. S., and Verhoogen, E. A. (2009). Exports and Wage Premia: Evidence from Mexican Employer-Employee Data. Unpub. paper, Columbia University.

Helpman, E., Itskhoki, O., and Redding, S. (2010). Inequality and Unemployment in a Global Economy. *Econometrica*, 78(4):1239–1283.

Hummels, D., J?rgensen, R., Munch, J., and Xiang, C. (2014). The Wage Effects of Offshoring: Evidence from Danish Matched Worker-Firm Data. *American Economic Review*, 104(6):1597–1629.

Irarrazabal, A., Moxnes, A., and Ulltveit-Moe, K. H. (2013). Heterogeneous Firms or Heterogeneous Workers? Implications for Exporter Premiums and the Gains from Trade. *The Review of Economics and Statistics*, 95(3):839–849.

Jann, B. (2008). The Blinder Oaxaca decomposition for linear regression models. *Stata Journal*, 8(4):453–479.

Krishna, P., Poole, J. P., and Senses, M. Z. (2014). Wage Effects of Trade Reform with Endogenous Worker Mobility. *Journal of International Economics*, 93(2):239–252.

Lentz, R., Piyapromdee, S., and Robin, J. M. (2016). The Anatomy of the Wage Distribution: How do Gender and Immigration Matter? 2016 Meeting Papers 1686, Society for Economic Dynamics.

Lund, C. G. and Vejlin, R. (2016). Documenting and Improving the Hourly Wage Measure in the Danish IDA Database. *Nationaløkonomisk Tidsskrift*, 1:1–35.

Macis, M. and Schivardi, F. (2016). Exports and Wages: Rent Sharing, Workforce Composition, or Returns to Skills? *Journal of Labor Economics*, 34(4):945–978.

Melitz, M. J. (2003). The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity. *Econometrica*, 71(6):1695–1725.

Munch, J. R. and Skaksen, J. R. (2008). Human capital and wages in exporting firms. *Journal of International Economics*, 75(2):363–372.

Oaxaca, R. (1973). Male-Female Wage Differentials in Urban Labor Markets. *International Economic Review*, 14(3):693–709.

Schank, T., Schnabel, C., and Wagner, J. (2007). Do exporters really pay higher wages? First evidence from German linked employer-employee data. *Journal of International Economics*, 72(1):52–74.

Woodcock, S. D. (2008). Wage differentials in the presence of unobserved worker, firm, and match heterogeneity. *Labour Economics*, 15(4):771–793.

Yeaple, S. R. (2005). A simple model of firm heterogeneity, international trade, and wages. *Journal of International Economics*, 65(1):1–20.

Appendix

A Additive model:

In this appendix we provide further details related to our analysis of the additive model (i.e. Sections 3, 4 and 4). First, in Section A.1 we provide more detail in relation to our estimation procedure. Next we explain how worker and firm fixed effects can be related to time-invariant observables and we show the results of these estimations (some of these results were also reported in the main text). Finally we outline the estimation procedure which is used in Section 4 and show how the difference in comparative advantage (difference in worker fixed effect under export and non-export) varies across industries.

A.1 Estimation Procedure

In this section we explain the estimation procedure for the baseline model and the extension to different returns across export status (this amounts to the models presented in Section 3). Our estimation procedure is as follows:

1. Estimate $\hat{\beta}$ and $\hat{\mu}_{ij}$ from $y_{it} = \mu_{i,j(i,t)} + x_{it}\beta + \epsilon_{ijt}$.
2. Decompose $\hat{\mu}_{ij} + \hat{\epsilon}_{ijt}$ into $\hat{\psi}_j$ and $\hat{\theta}_i$ from $\hat{\mu}_{ij} + \hat{\epsilon}_{ijt} = \psi_j + \theta_i + \alpha_{ij} + \xi_{ijt}$, assuming $E\left[\xi_{ijt}|\psi_j, \theta_i, \alpha_{ij}\right] = 0$ and $E\left[\alpha_{ij}|\psi_j, \theta_i\right] = 0$.

In the extension with export-specific returns (see Section 3.1.2) step 1 is estimated separately for the export and non-export samples.

In the estimation procedure we use the STATA routine reghdfe developed by Sergio Correia (Correia, 2016). The routine solves high-dimensional fixed effects problems with multi-way fixed effects. The reghdfe does not solve the least-square problem by solving the normal equations directly. Instead it uses the Method of Alternating Projections to solve the model (see Correia (2016)).

For step 1, the main reason for using reghdfe is that it standardizes the expectation of each set of fixed effects to zero (i.e. it takes out the constant). In our case, the first set of fixed effects is the job spell fixed effects and the second set of fixed effects is the year effects. For these two sets of fixed effects we impose that the sample weighted sum of each individual set of fixed effects is 0 (in reghdfe this amounts to "absorbing" these fixed effects). This normalization is needed to solve the "omitted group problem" explained in Jann (2008) and Fortin et al. (2011). Thus, in case of export-specific returns, the Oaxaca-Blinder decomposition is robust to the choice of reference (or omitted) group. Thus, it no longer matters for the decomposition which year dummy we omit. We add the constant from the regressions within export and non-export to the match effects $\hat{\mu}_{ij}$. Additionally, we add the normalized year dummies to $\hat{\beta}$.

Step 2 is again solved using the reghdfe routine. This time we absorb the worker and firm fixed effects. Note, that they too are normalized such that the sample weighted sum of each individual set of fixed effects is 0. Note also, firms have trade-specific firm fixed effects. An arbitrary firm effect is omitted in estimation. The constant from this regression is treated as a separate entry which cancels out once we look at the difference between exporters and non-exporters.

A.2 Decomposing the worker and firm effects

Following Woodcock (2008) we investigate how much of the variation in worker and firm effects can be explained by observable time-invariant variables. In particular, we clean the worker effect of the effect of the interaction of gender and education. Similarly, we clean the firm effects of the effect of industry. These characteristics are time-invariant, and hence not included in the main estimation. However, under the assumption of orthogonality from any cleaned worker and firm effects, we can use variation in the cross-section to estimate the effect of these time-invariant characteristics. To formalize the above, assume that the firm effect consists of an observable part $x_j^{C,F}$ and a true unobservable fixed effect v_j, then

$$\psi_j = x_j^{C,F}\beta^{C,F} + v_j + \xi_j$$

where we assume $E[\xi_j|v_j, x_j^{C,F}]$=0 and $E[v_j|x_j^{C,F}] = 0$. And the for the worker fixed effect

$$\theta_i = x_i^{C,W}\beta^{C,W} + u_i + \delta_i$$

where $E[\delta_i|u_i, x_i^{C,W}] = 0$ and $E[u_i|x_i^{C,W}] = 0$.

We use 16 industry groupings corresponding to two-digit NACE codes, with some categories merged to get a consistent mapping backward in time. Recall that 120 percent of the wage gap was due to differences in firm effects (see Table 4). When we use these industry categories it accounts for 18 percent of the firm fixed effects gap. Put differently 22 percent of the exporter wage gap is accounted for by industry. This still leaves 97 percent to pure unobserved firm characteristics. Our results suggest that there are both between firm within industry and between industry differences, but within differences are more important.

Worker fixed effects contributed negatively to the total exporter wage gap by -18 percent (-0.006). Interestingly, the decomposition of worker fixed effects into gender-education-specific effects suggests that these characteristics would increase the total wage gap by 0.003, meaning that worker unobservables actually contribute -0.008 or -27 percent. This strengthens the point made earlier that controlling for worker unobservables is important, when testing theories of firm self-selection into export.

A.3 Comparative advantage

A.3.1 The normalization procedure

We proceed as follows:

1. We select all firms with valid value-added data. We trim the data at the 1st and 99th percentiles of the value added per worker distribution. Next, we create a set of firms selecting the largest set of connected firms in the export and non-export sector, respectively. We refer to this set as the dual-connected set.

2. To account for nominal trends in value-added (and potentially smooth out any measurement error), we estimate a firm-level fixed effects model of log-average value-added per worker on year dummies. Firms are represented by their value-added fixed effect, $\bar{va}_j$.

3. We re-estimate the model of Section 3.1.2 on the dual-connected set applying the usual normalization in an AKM model. We run two regressions, one for the exporters and one for non-exporters. At this stage the fixed effects from the two models are not comparable. Denote the firm fixed effects from this regression $\hat{\psi}_j^h$

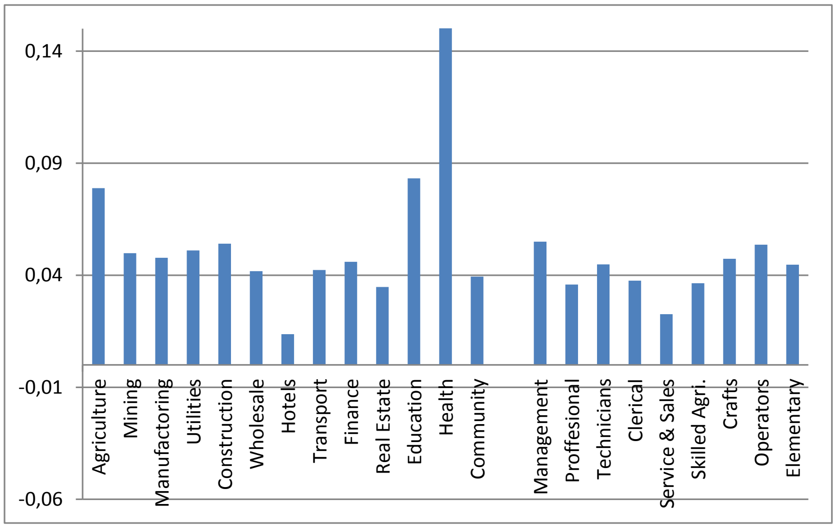
4. We then follow Card et al. (2016) and run a series of regressions using the firm fixed effects from step 3, $\hat{\psi}_j^h$, on export status dummies (c^h) and export status specific returns to the value-added fixed effects, $\bar{va}_j$, where firm fixed effects under a varying threshold τ are set to zero (importantly the threshold is independent of export status). That is, for $h = \{E, NE\}$ we estimate

$$\hat{\psi}_j^h = c^h + \gamma^h \max\{\bar{va}_j - \tau, 0\}.$$

varying τ. c^h is the intercept of the regression for $h = \{E, NE\}$, while γ^h is the slope. τ is the value under which value added per worker does not influence the wage the worker receives. We select the τ that minimizes the root mean squared error. Note that the regressions are employment weighted such that an observation is a worker-year.

   (a) We also run a version of the regression, where we allow for a quadratic and a cubic term in $\max\{\bar{va}_j - \tau, 0\}$ to get a better fit to data. This does not alter the assumption, i.e. firms below a certain threshold have no surplus for its workers to bargain over. The more flexible form is only

Figure 3: Comparative Advantage by Industry and Occupation



The figure shows the average worker fixed effect gap ($\theta_i^E - \theta_i^{NE}$) for NACE one digit industries and ISCO one digit occupations.

used to get a more credible kink-point (i.e. the last firm for which the firm fixed effect is the same under export and non-export).

5. Finally we re-estimate the model of Section 3.1.2 on the dual-connected set. We run two regressions, one for the exporters and one for non-exporters. The regressions apply the usual normalization in an AKM model but we now make sure that the "normalizing firm" is now the set of firms (we treat the set of firms as one firm) that have $\bar{va}_j < \tau$ in step 4. This last step essentially anchors the fixed effects from the two samples and makes them comparable. The assumption essentially implies that the firm effect for the normalizing firm does not change with export status. We do not require that the firm effect is 0.

   (a) The change in worker effects going from export to non-export is now our measure of comparative advantage. We can only quantify this for workers who appear in both the export and non-export sample. Identification of this object comes from the fact that we observe these workers making job moves from/to the normalizing firm to/from all the other firms. The "normalizing firm" serves as the unit that connects the two samples (export and non-export). Given a connected set of workers and firms the usual identification arguments for separating worker and firm fixed effects apply.

A.3.2 Supplementary Results

Online Appendix - supplementary online material

B Data

This part of the appendix provides further details on the value-added data we use for the analysis. We also provide more information about our sample selection procedure. Finally we repeat our baseline decompositions for different definitions of export status.

B.1 GF

The VA sample and BLM sample introduced in the main text consist of firms which can be found in the General Firmastistik (GF) database. GF contains firm level accounting data and value added tax (VAT) records. GF is only available for a selected subset of Danish firms, i.e. those who are sufficiently active. The activity requirement excludes firms that have mandatory pension contributions equivalent to less than half a year of full-time employment and firms with low earnings. In the sample period roughly 15-20 percent of worker observations have missing productivity measures each year before evoking the connectivity criterion. The equivalent number for firms is roughly 25-30 percent each year. Thus, we only have information on productivity (value added) for a sub-sample of the universe of Danish firms, i.e. the larger firms.

B.2 Sample Selection

This section describes the sample selection and provides basic description of the sample. As explained in Section 2.1 we focus on the period 2004-2010 to ensure a consistent definition of export on the extensive margin. Firms are defined as exporters if they have positive outside-EU export within a given year. Table 9 gives the number of workers and firms for the total sample. First, we discard jobs that are not considered main employment. Second, we discard all public sector jobs. Third, we discard individuals with no or poor quality educational information. Fourth, we define labor market to be the time of completion for the highest education obtained in 2010 and discard all observations before labor market entry or where the worker is younger than 19. Fifth, we delete observations with no or invalid firm identifiers or where the hour wage measure is of bad quality. Finally, we trim the 1 and 99 percentiles in the wage distribution each year in order to mitigate the effect of outliers. Since we will estimate a model that is similar of nature to the AKM model, we focus on the largest set of connected firms as described by Abowd et al. (2002). Our final sample has 7.355.114 person-years. 20.3 percent of all firms are dropped when considering the connected set. However, only 3.4 percent of all person-years are dropped, so the firms in the connected set still account for almost all employment and aggregate export.

Table 9: Sample Selection

	Workers	Firms
Initial panel	4.662.110	375.667
Drop nonemployed	3.400.660	233.784
Drop public sector	2.514.424	223.374
Delete individuals older than 55	2.514.424	223.374
Delete workers who has only missing education info.	2.420.700	221.673
Delete obs. pertaining to spells before labor market entry	2.420.700	218.770
Delete obs. before the year where the worker turns 19	2.218.450	215.346
Disregard workers who are observed under education after completion of the highest completed education	1.955.133	210.860
Disregard ind. with age-education length<5	1.955.097	210.857
Disregard ind. with entry year before finished education	1.947.426	210.541
Disregard obs. with an invalid cvrnr (firm identifier)	1.936.469	210.540
Disregard low quality wage observations	1.767.297	189.866
Trim wages	1.744.186	186.828
Largest connected set	1.685.179	148.868
Percent lost due to connectivity	3.4 %	20.3 %

The table shows the results of the sample selection procedure. Each row reports the remaining number of observations after an additional sample restrictions have been invoked.

Table 10: Decompositions with Alternative Export Definitions

	Baseline			Export-Specific		
	(1)	(2)	(3)	(4)	(5)	(6)
	>0	>100,000	>500,000	>0	>100,000	>500,000
Raw Gap	0.031	0.028	0.03	0.031	0.028	0.03
Observables	0.007	0.007	0.011	-0.001	0.002	0.007
Firm FE	0.03	0.03	0.032	0.038	0.036	0.037
Worker FE	-0.005	-0.009	-0.013	-0.006	-0.009	-0.014
Match FE	0.00	0.00	0.00	0.00	0.00	0.00

The table replicates the decomposition of Section 3 with alternative definitions of an exporter.

B.3 Export Status

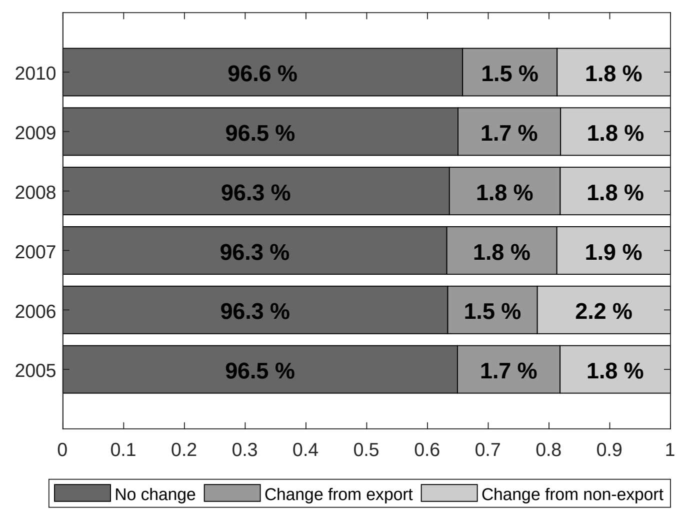
In this section we describe patterns on the extensive margin of export status, i.e. how many firms and workers change export status and in which direction? Figure 4 shows that roughly 2 percent of all firms (or 3 percent of all workers) change export status in each direction each year. Firms starting to export are somewhat larger than those that stop exporting.

B.4 Changing the definition of export

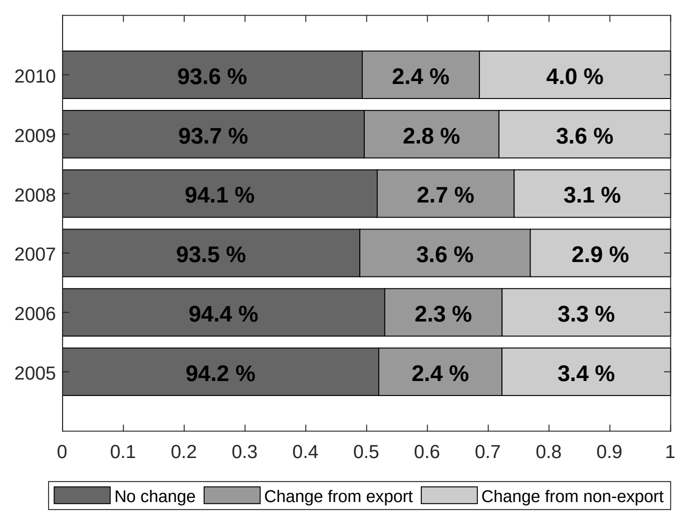
In our analysis we have defined an exporter in a certain year as any firm with any export within that year. However, this means that firms sending anything outside EU borders will be termed a exporter on the same terms as large firms with large exports. To check whether our results are affected by these marginal and potentially noisy exporters, we increase the minimal requirement for being an exporter from 0 to 100.000 DKK per year and 500.000 DKK per year, respectively. The results from Section 3 are reproduced with the alternative export definitions in Table 10. Columns (1) to (3) compare the baseline decomposition, while Columns (4) to (6) compare the decomposition for the model with export-specific returns to observable characteristics. The decompositions are very robust across the different definitions of export status.

Figure 4: Firms changing export status.

(a) Firm level



(b) Worker level



Panel (a) shows the number of firms changing export status each year . Panel (b) shows the same number weighted by the sample firm size.

C Non-additive model

This appendix provides some more details about the decompositions we use in the BLM model presented in the main text.

C.1 Version 1: Standard

The normalizations are:

- we normalize the interaction effects τ_{khl} such that for each $k = \{1, .., K\}$ we have $E[\tau_{kh\alpha}|k, h] = 0$ for $h = \{E, NE\}$ and for each $l = \{1, .., L\}$ we have $E[\tau_{khl}|l] = 0$.
- in addition we assume that workers do not have export-specific returns, i.e. $\theta_{lE} = \theta_{lNE}$.

Given these normalizations presented in the main text we can write the expected wage conditional on export status $h = \{E, NE\}$ as

$$\begin{aligned} E[w_{it}|h] &= \sum_{k=1}^{K} P[k_{it} = k|h]\, \psi_{kh} + \sum_{l=1}^{L} P[l_i = l|h]\, \theta_l \\ &+ \sum_{k=1}^{K} P[k_{it} = k|h] \sum_{l=1}^{L} P[l_i = l|k_{it} = k, h]\, \tau_{khl} \end{aligned}$$

Thus, the exporter wage gap is composed of

$$\begin{aligned} & E[w_{it}|E] - E[w_{it}|NE] \overset{u.e}{=} \\ \text{1. firm composition} \quad & \left[\sum_{k=1}^{K} \{P[k_{it} = k|E] - P[k_{it} = k|NE]\}\, \psi_{kE}\right] \\ \text{2. differential firm returns} \quad + & \left[\sum_{k=1}^{K} P[k_{it} = k|NE]\, [\psi_{kE} - \psi_{kNE}]\right] \\ \text{3. worker composition} \quad + & \left[\sum_{l=1}^{L} \{P[l_i = l|E] - P[l_i = l|NE]\}\, \theta_l\right] \\ \text{4. match composition} \quad + & \left[\sum_{k=1}^{K} \sum_{l=1}^{L} \{P[k_{it} = k|E]\, P[l_i = l|k_{it} = k, E]\right. \\ & \left. - P[k_{it} = k|NE]\, P[l_i = l|k_{it} = k, NE]\}\, \tau_{klE}\right] \\ \text{5. differential match returns} \quad + & \left[\sum_{k=1}^{K} \sum_{l=1}^{L} P[k_{it} = k|NE]\, P[l_i = l|k_{it} = k, NE]\right. \\ & \left. \cdot \{\tau_{klE} - \tau_{klNE}\}\right] \end{aligned}$$

where $u.e.$ refers to the fact that the shown decomposition is carried out under the returns "under export". Similarly to the Oaxaca-Blinder decomposition, a decomposition under the returns "under non-export" also exists. We interpret the first term as differences in firm composition, the second term is the effect of firm wage policy ("returns"), the third term as differences in worker composition, the fourth term is the differences in the composition of match specific effects, while the fifth term is differences in the size of match specific effects. Thus, the exporter wage gap is explained by differences in composition and return of firms and matches, and differences in worker composition. We add to the baseline decomposition, by allowing exporting firms to form more beneficial matches.

Note that $P[k_{it} = k|h]$ can be recovered after the firm classification step, while $P[l_i = l|k_{it} = k, h]$ can be recovered from the estimated model.

C.2 Version 2: Card-Style

In this section, we allow for export-specific worker effects as in Section 4. Intuitively, we treat worker, firm and match effects as if they came from two separate BLM models. In this spirit, we assume that for each $h = \{E, NE\}$ we have $E[\tau_{khl}|k, h] = 0$ for each k and we have $E[\tau_{khl}|l, h] = 0$ for each l. Thus, we interpret match effects as deviations from the export-specific combination of worker and firm effects. To separate worker and firm effects we need one normalization within each sector. Here, we will assume that the lowest firm class has so low productivity that their firm effect is zero both when exporting and not (therefore the first two summations for $k = 1$ are not included below). This mirrors the normalizing assumption of Section 4. As in Section 4 this is an economic assumption that influences the decomposition. The normalization allows us to do a full decomposition:

$$
\begin{aligned}
&E[w_{it}|E] - E[w_{it}|NE] \overset{u.e}{=} \\
&\text{1. firm composition} \quad \left[\sum_{k=2}^{K} \{P[k_{it} = k|E] - P[k_{it} = k|NE]\}\, \psi_{kE}\right] \\
&\text{2. differential firm returns} \; + \left[\sum_{k=2}^{K} P[k_{it} = k|NE]\, [\psi_{kE} - \psi_{kNE}]\right] \\
&\text{3. worker composition} \; + \left[\sum_{l=1}^{L} \{P[l_i = l|E] - P[l_i = l|NE]\}\, \theta_{lE}\right] \\
&\text{4. differential worker returns} \; + \left[\sum_{l=1}^{L} P[l_i = l|NE]\, \{\theta_{lE} - \theta_{lNE}\}\right] \\
&\text{5. match composition} \; + \left[\sum_{k=1}^{K}\sum_{l=1}^{L} \{P[k_{it} = k|E]\, P[l_i = l|k_{it} = k, E]\right. \\
&\qquad \left. - P[k_{it} = k|NE]\, P[l_i = l|k_{it} = k, NE]\}\, \tau_{klE}\right] \\
&\text{6. differential match returns} \; + \left[\sum_{k=1}^{K}\sum_{l=1}^{L} P[k_{it} = k|NE]\, P[l_i = l|k_{it} = k, NE]\right. \\
&\qquad \left. \cdot \{\tau_{klE} - \tau_{klNE}\}\right]
\end{aligned}
$$

where *u.e.* refers to the fact that the shown decomposition is done under the returns "under export". Similarly to the Oaxaca-Blinder decomposition, a decomposition under the returns "under non export" also exists. We interpret the first term as differences in firm composition, the second term is the effect of firm wage policy ("returns"), the third term as differences in worker composition, the fourth term is differences in return export-specific worker skills, the fifth term is the differences in the composition of match specific effects, while the sixth term is differences in the size of match specific effects. Thus, the exporter wage gap is explained by differences in composition and return of worker, firms and matches.

Table 11: Decomposition of the Exporter Wage Gap with Non-Additive Wages under Non-Exporter Returns

2009-2010	**Standard (1)**	**- AKM (2)**	**CCK-type (3)**	**- AKM (4)**
		under non-exporter returns		
Total	0.046	0.046	0.046	0.046
Worker	0.015	0.003	0.044	0.040
- Composition			*0.014*	*0.002*
- Return			*0.029*	*0.036*
Firm	0.031	0.028	0.002	-0.009
- Composition	*0.025*	*-0.004*	*0.026*	*0.005*
- Return	*0.006*	0.003	*-0.024*	*-0.042*
Match	0.000	0.000	0.000	0.000
- Composition	*-0.002*	*-0.001*	*-0.003*	*-0.001*
- Return	*0.002*	*0.000*	*0.003*	*0.000*

The table shows the decomposition of the exporter wage gap from the BLM model estimated on the years 2009-2010 under the assumptions of Section 5.1. All decompositions are under the non-exporter returns (the similar decomposition under exporter return is presented in the main text, see Table 8). Column (1) is based on the restrictions N1 (see Section 5.1), while Column (2) shows the AKM equivalent. Column (3) is based on N2, while Column (4) shows the AKM equivalent. For the AKM decompositions (Column 2 and 4) some elements of the decomposition are omitted and the detailed decomposition of worker, firm and match effects is only based on the appropriate identifying sub-samples, therefore rows do not sum the total.

Table 12: Decomposition of the exporter wage gap with Non-Additive Wages

$K = L = 2$	**Standard (1)**	**CCK-type (2)**	$K = L = 10$	**Standard (1)**	**CCK-type (2)**
	under export returns			*under export returns*	
Total	0.046	0.046	**Total**	0.046	0.046
Worker	0.021	0.032	**Worker**	0.010	0.030
- Composition		*0.022*	*- Composition*		*0.010*
- Return		*0.010*	*- Return*		*0.020*
Firm	0.025	0.014	**Firm**	0.036	0.016
- Composition	*0.027*	*0.027*	*- Composition*	*0.025*	*0.026*
- Return	*-0.002*	*-0.013*	*- Return*	*0.011*	*-0.010*
Match	0.000	-0.000	**Match**	-0.000	-0.000
- Composition	*0.000*	*0.000*	*- Composition*	*-0.002*	*-0.000*
- Return	*-0.000*	*-0.000*	*- Return*	*0.002*	*0.000*

The table shows the decomposition of the exporter wage gap from the BLM model estimated on the years 2009-2010 under the assumptions of Section 5.1 and with either 2 or 10 worker and firm types. Column (1) is based on the restrictions N1 and Column (2) is based on N2. All decompositions are under the exporter returns (the similar decomposition under non-exporter return delivers very similar results).

D Additional Tables

Table 11 shows the same decompositions as Table 8 but under the returns of non-exporters.

Table 12 shows the BLM style decomposition from Table 8, i.e. Column (1) and (3), using a smaller number of firm classes and worker types. The conclusions are very robust.

Table 13: Decomposition of the Exporter Wage Gap with Non-Additive Wages under Exporter Returns

2006-2007	**Standard (1)**	**CCK-type (3)**
	under exporter returns	
Total	0.021	0.021
Worker	0.000	0.014
Firm	0.021	0.007
Match	0.000	0.000

The table shows the decomposition of the exporter wage gap from the BLM model estimated on the years 2006-2007 under the assumptions of Section 5.1. Column (1) is based on the restrictions N1, while Column (3) is based on N2. All decompositions are under the exporter returns.