Offshoring and Firm Overlap*

Stella Capuano
IAB Nuremberg
University of Hagen

Hartmut Egger†
University of Bayreuth
CESifo, GEP, and IfW

Michael Koch
University of Bayreuth

Hans-Jörg Schmerer
University of Hagen
CESifo and IAB

June 17, 2016

Abstract

We set up a model of offshoring with heterogeneous producers that captures two empirical regularities of German offshoring firms. There is selection of larger, more productive firms into offshoring. However, the selection is not sharp, and offshoring and non-offshoring firms coexist over a wide range of the revenue distribution. An overlap of offshoring and non-offshoring firms emerges in our model because, in contrast to textbook models of trade with heterogeneous producers, we allow firms to differ in two technology parameters thereby decoupling the offshoring status of a firms from its revenues. In an empirical exercise, we employ firm-level data from Germany to estimate key parameters of the model and show that ignoring the overlap lowers the estimated gains from offshoring by more than 67 percent and exaggerates the importance of the extensive margin for explaining the evolution of German offshoring over the last 25 years.

JEL-Classification: C30, F12, F14

Keywords: Offshoring, Heterogeneous firms, Firm overlap, Quantitative trade model, Extensive and intensive margin of offshoring

*We thank Holger Breinlich, Swati Dhingra, Benno Heid, Mario Larch, John Morrow, Peter Neary, and Veronica Rappoport for encouraging discussion and useful suggestions on earlier versions of this paper. We are grateful to participants at the Annual Meetings of the European Economic Association, the European Trade Study Group, the International Association for Applied Econometrics, and the German Economic Association, the Danish Trade Workshop, the Dynamics, Economic Growth, and International Trade Conference, the World Congress of the Econometric Society, and the Workshop in the Honor of Josef Falkinger as well as research seminars at the Chinese University of Hong Kong, London School of Economics, Shanghai University of Finance and Economics, University of Aarhus, University of Bergen, University of Bielefeld, University of Duesseldorf, University of Duisburg-Essen, University of Essex, University of Hagen, University of Heidelberg, University of Trier, University of Tuebingen, University of Uppsala, University of Wuerzburg, and University of Xiamen for helpful comments.

†Corresponding author: University of Bayreuth, Department of Economics, Universitätsstr. 30, 95447 Bayreuth, Germany; Email: hartmut.egger@uni-bayreuth.de.

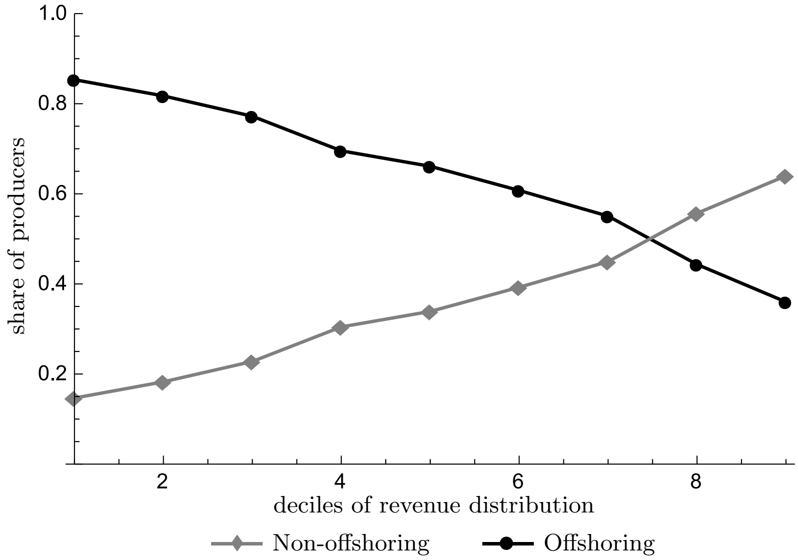
1 Introduction

Offshoring and its effects on domestic labor markets have played a prominent role in academic research and the public debate over the last two decades. In recent years, attention has shifted towards understanding the specific nature of firms that choose to offshore. Relying on models of heterogeneous firms, trade economists have pointed out that similar to exporters, offshoring firms are exceptional producers. They are larger, more productive, and make higher profits than their purely domestic competitors (see Antràs and Helpman, 2004; Antràs et al., 2006; Egger et al., 2015). Although this pattern finds strong support in the data (cf. Bernard et al., 2012; Hummels et al., 2014; Moser et al., 2015), existing theoretical work misses an important empirical fact that is illustrated for German producers in Figure 1. There is a considerable overlap of offshoring and non-offshoring firms in the data, because some but not all firms from the various deciles of the revenue distribution choose to offshore.[1]

To explain this fact and to shed light on how it changes the conclusions we draw when it comes to the consequences of offshoring are the aim of this paper. For this purpose, we construct a theoretical model that captures two features that are salient for the empirical pattern of offshoring. *Selection*, because offshoring is more common among producers from higher quantiles of the revenues distribution; *Overlap*, since there is coexistence of offshoring and non-offshoring producers over a wide range of the revenue distribution. After a thorough theoretical analysis, we structurally estimate key parameters of our model, using firm-level data from Germany. Based on these parameter estimates, we then study the nature and extent of the bias in the quantitative welfare effects of offshoring that originates from disregarding the overlap in the data and show how ignoring this overlap affects the relative importance of extensive and intensive margins for explaining observed changes of offshoring.

In the theory section, we set up a stylized two-country model of offshoring, with labor being the only factor of production. The two countries differ in their level of development and since offshoring is low-cost seeking, it is one directional and leads to production shifting from the more developed source country to the less developed host country. Following Acemoglu and Autor (2011), we model production as the assembly of tasks, with firms differing in the number of tasks

[1] In the appendix, we provide further evidence on the overlap of offshoring and non-offshoring firms, using domestic employment or the number of tasks performed for categorizing firms. We also report the overlap of offshoring and non-offshoring firms for manufacturing and service industries. The main conclusion from Figure 1 is robust to these changes.



Source: IAB establishment panel, covering 20,334 establishment observations for Germany in the years 1998, 2000, 2002 from all industries and size categories; Descriptive statistics rely on own computations, using inverse probability weights.

Figure 1: Share of offshoring firms in different revenue categories

performed in the production process. The number of tasks is directly linked to firm productivity, reflecting the idea that more tasks allow for a stronger division of labor in the production of goods.[2] Hence, firm heterogeneity is rooted in differences in the task production process, which are exogenous to the firm. Since one source of heterogeneity is not sufficient to model overlap in a setting that features selection of firms into offshoring, we assume that firms also differ in the share of tasks that can be offshored to the low-wage host country.[3] In the tradition of theoretical work building on the Melitz (2003) framework, we model firm heterogeneity as the outcome of a lottery, but acknowledge that firms draw *two* technology parameters: the number of tasks and the share of offshorable tasks.[4]

[2]Whereas we cannot observe firm productivity in our dataset, we can find indirect support for this assumption when looking at the number of tasks and the log revenues of non-offshoring firms, which are the firms for whom we have full information on the task content of production. The correlation coefficient between these two variables is positive and reaches a high value of 0.74 in our establishment sample.

[3]Becker et al. (2013) point out that in order to be offshorable, a task must be routine (cf. Levy and Murnane, 2004) and lack the necessity of face-to-face contact (cf. Blinder, 2006). Blinder and Krueger (2013) classify only 25 percent of US jobs as being vulnerable to offshoring according to these criteria.

[4]It is well established that allowing for firm heterogeneity in more than just one dimension helps making the Melitz (2003) model better suited for explaining firm-level evidence in the context of trade. Prominent examples that provide extensions in this direction include Davis and Harrigan (2011), Hallak and Sivadasan (2013), Armenter and Koren (2015), and Harrigan and Reshef (2015). Recent empirical applications have accounted for various sources of heterogeneity to achieve a better fit with the data (cf. Das et al., 2007; Eaton et al., 2011;

It is the interaction of the two technology parameters that determines the offshoring decision in our model. It is possible that a firm operating a sophisticated technology with many different tasks finds itself in a position with none of its tasks being offshorable, despite its high productivity and large revenues. However, it is also possible that a firm with a simple technology requiring only a few tasks can offshore a significant share of these tasks. This establishes a technology-driven overlap, which itself does not depend on the choice of firms and is thus exogenous. To give firms a role in our model, we assume that offshoring is subject to a fixed cost, and hence the gains from offshoring must be sufficiently high to make it attractive for firms. This makes selection into offshoring an important feature of our model and points to changes in the share of offshoring firms as a key determinant of overlap in this setting. Provided that the fixed costs of offshoring are sufficiently large, offshoring in our model is more attractive ceteris paribus for firms operating a technology with more tasks and thus featuring higher revenues, which is well in line with the evidence reported in Figure 1.

We use this framework to analyze how changes in variable and fixed offshoring costs affect offshoring and welfare in the source country. A decline in the variable cost of offshoring lowers the price of foreign workers. This makes offshoring attractive for a wider range of producers and increases the volume of tasks imported by incumbent offshoring firms – because the cost of importing tasks performed abroad makes them more competitive and because they substitute domestically produced tasks for imported ones. Both effects stimulate labor demand in the host country and lead to a rise in foreign wages. However, the increase in foreign wages is of second order and dominated by the initial drop in variable offshoring costs, so that the effective cost of employing foreign workers decreases. This reflects an appreciation of domestic relative to foreign labor and thus an improvement of the (double) *factorial* terms of trade for the source country of offshoring with positive welfare implications (cf. Ghironi and Melitz, 2005). Things are different if the fixed cost of offshoring falls. Whereas this makes offshoring attractive for new producers, the higher foreign labor demand and the resulting increase in host country wages prompt incumbent offshoring producers to reduce the volume of imported tasks. The deterioration of the (double) factorial terms of trade counteracts the direct welfare gain from a lower offshoring fixed cost and it can lead to the somewhat counterintuitive result that lifting a technology barrier can actually lower welfare of the source country of offshoring.[5]

Helpman et al., 2016).

[5]Relying on the relative effective labor costs when providing intuition for the welfare effects of offshoring

In the second part of the paper, we combine three datasets from two different sources. The first one is the Establishments Panel of the Institute for Employment Research (IAB), which provides information on a (representative) sample of 16,000 establishments of all branches of the economy and all size categories from annual surveys. In 1999, 2000, and 2003, the dataset reports offshoring information of German establishments from the previous business year. In addition, we use the BIBB-BAuA 2006 "Survey of the Working Population on Qualification and Working Conditions in Germany" from the Federal Institute for Occupational Safety and Health to construct a measure of task content for 341 occupations. We finally use the Linked Employer-Employee Database from the Institute for Employment Research (LIAB) to aggregate the task composition at the occupation level to the establishment level. This gives a unique dataset for studying offshoring in the context of task production, and we use this dataset to estimate key parameters of our theoretical model. For this purpose, we employ a methods of moments estimation strategy that minimizes the distance between selected moments from the data and the model. More specifically, we employ the minimum distance estimator of Ferguson (1958), since it is particularly suited for our application (see Cameron and Trivedi, 2005, for a discussion of this approach). With the parameter estimates at hand, we assess the fit of the quantitative model with targeted and not targeted data moments. Overall, our model somewhat underestimates the overlap and the mean of marginal production costs, but does a good job in capturing the variance of marginal costs and the link between marginal costs and revenues in the data.

We also analyze to what extent ignoring the overlap in the data matters for the conclusions drawn from the model. To tackle this issue, we construct an otherwise identical model with sharp selection into offshoring and thus no overlap, and estimate the parameters of this alternative model variant, using the methods of moments approach outlined above. We show that ignoring the overlap in the data significantly lowers the estimated cost saving from offshoring. This is intuitive, because the model without overlap presumes that all firms that make use of offshoring are high-productivity producers and these firms require a lower cost saving to find offshoring attractive. The discrepancy between the two models regarding the estimated cost savings from offshoring generates quantitatively sizable differences in the welfare effects attributed to offshoring by these models. The model with overlap associates the observed share

acknowledges that trade involves the exchange of final against intermediate goods, so that changes in the relative price of exports and imports goods do not reflect changes in the terms of trade in consumer goods. We show in the appendix that a worsening of the factorial terms of trade for the source country is instrumental for the existence of welfare losses from offshoring in the source country.

of offshoring establishments with a welfare gain for Germany that amounts to an increase in GDP per capita of 15.57 percent. The welfare gain attributed to offshoring is reduced to 5.12 percent in the model variant without overlap, which is more than 67 percent lower than in the model with overlap.

We finally use our quantitative model to decompose the observed increase of German offshoring openness from 18.03 percent in 1990 to 30.26 percent in 2014 into its *intensive* margin – capturing changes in the offshoring activity of incumbent offshoring firms – and its *extensive* margin – capturing changes in the mass of offshoring firms.[6] We show that both margins contributed significantly to the observed increase of German offshoring, with the extensive margin gaining importance in later years. Ignoring the observed overlap of offshoring and non-offshoring establishments leaves only little scope for the intensive margin and leads to a severe exaggeration of the importance of the extensive margin in explaining the evolution of German offshoring openness over the period 1990-2014. The model with overlap suggests that this increase in offshoring openness has increased German welfare by 4.88 percent and has contributed 12.67 percent to the overall increase in German GDP per capita over this period.

Shedding light on the overlap of offshoring and non-offshoring firms in the revenue distribution, our analysis is most closely related to Armenter and Koren (2015), who have documented an overlap for exporting and non-exporting firms, using US census data. To reconcile the predictions of the Melitz (2003) model regarding the composition of exporters with the data, Armenter and Koren suggest randomizing fixed exporting costs to add an additional source of heterogeneity.[7] Contrasting the thus extended with the original Melitz model, they find that a model with sharp selection into exporting significantly overestimates the role of entry and exit into the export market for the growth of exports. Our analysis differs from Armenter and Koren (2015) in several important dimensions. First, we document and analyze the overlap of offshoring and non-offshoring firms instead of exporting and non-exporting firms. Second, we allow for dependencies in the distributions of technology parameters, and show that such dependencies are important for describing the overlap in the data. Third, we provide a detailed welfare analysis

[6] We measure offshoring openness by the import of intermediate goods and services relative to GDP.

[7] Whereas we could also consider the the fixed cost of offshoring as second source of firm heterogeneity, we do not follow this path because it is a particularly attractive feature of our approach that we can aggregate both sources of heterogeneity into differences of marginal costs, which then subsume heterogeneity of firms in all relevant performance measures. Accordingly, once the distribution of marginal costs has been specified, we can rely on the toolbox of heterogeneous firms models along the lines of Melitz (2003) for our analysis.

and show that ignoring the overlap in the data leads to a severe downward bias in the welfare effects predicted by our quantitative trade model. Fourth, we offer a rigorous decomposition analysis for the observed changes in German offshoring openness over the last 25 years and point to a particularly strong exaggeration of the extensive margin in the context of offshoring.[8]

By studying the effects of offshoring, our model contributes to a large body of literature that includes prominent contributions by Grossman and Rossi-Hansberg (2008), Rodriguez-Clare (2010), and more recently Acemoglu et al. (2015). Thereby, we associate offshoring with a relocation of some tasks for production in a low-wage country, as in Grossman and Rossi-Hansberg (2008). However, assuming that offshoring is subject to constant per unit costs that are the same for all tasks, firms, when choosing to offshore at all, shift production of all (technologically) offshorable tasks abroad. This closes endogenous adjustments in the number of offshored tasks as a channel through which firms can respond to changing offshoring costs. Focussing on the decision of heterogeneous firms to offshore, we follow Egger et al. (2015) and emphasize the adjustments in the mass and type of offshoring firms, whose quantitative importance has been put forward by recent empirical evidence (cf. Bergin et al., 2011).[9] In addition to adjustments at the *extensive* margins of offshoring, the assumption of a Cobb-Douglas production technology allows for an *intensive task margin*, which captures changes in the volume of imported task output for a given composition of firms and tasks.

The remainder of the paper is organized as follows. In Section 2 we set up the theoretical model, distinguish important adjustment margins, and study the welfare effects of offshoring in the presence of overlap. In Section 3 we describe the dataset, estimate key model parameters, discuss the goodness of fit of our model, and quantify the welfare effects of offshoring. Section 4 provides a discussion on how our results change, if we do not account for the overlap of offshoring and non-offshoring firms in the data. Section 5 applies our model to decompose the observed increase in German offshoring openness between 1990 and 2014 into its extensive and intensive margin and sheds light on the welfare gains attributable to the increase in offshoring over this period. In Section 6 we study the robustness of our results for subsamples of industries and host

[8]The only other study that explains as we do an overlap of offshoring and non-offshoring firms is Rodriguez-Lopez (2014), who formulates a probabilistic model of offshoring and shows that the interaction of a *selection* effect and an *escape-competition* effect produce a hump shape relationship between firm productivity and offshoring probability. Provided that revenues are positively correlated with productivity, our data does not support a hump shape in this relationship.

[9]The interaction of task and firm margins of offshoring and the consequences for the creation and destruction of domestic jobs has been recently studied by Egger et al. (2016).

countries of offshoring and analyze whether our results change when using a different estimation strategy. The last section concludes with a summary of the most important results.

2 A model of offshoring and firm overlap

2.1 Basic assumptions and intermediate results

We consider a world with two economies. Consumers in both countries have CES preferences over a continuum of differentiated and freely tradable goods $y(\omega)$. The representative consumer's utility is given by $U = \left[\int_{\omega\in\Omega} y(\omega)^{(\sigma-1)/\sigma} d\omega\right]^{\sigma/(\sigma-1)}$, where $\sigma > 1$ is the elasticity of substitution between different varieties ω and Ω is the set of available consumer goods. Maximizing U subject to the representative consumer's budget constraint $I = \int_{\omega\in\Omega} p(\omega)y(\omega)$ gives isoelastic demand for variety ω:

$$y(\omega) = \frac{I}{P}\left[\frac{p(\omega)}{P}\right]^{-\sigma}, \tag{1}$$

where I is aggregate income, $p(\omega)$ is the price of good ω, and $P = \left[\int_{\omega\in\Omega} p(\omega)^{1-\sigma} d\omega\right]^{1/(1-\sigma)}$ is a standard CES price index.

The two economies differ in their level of development and are populated by L and L^* units of labor, respectively, where an asterisk refers to the economy with the lower level of development. This is the host country of offshoring, whereas the more advanced economy is the source country of offshoring. Similar to Egger et al. (2015), we assume that the host country lacks the technology to operate its own firms. This implies that all (industrial) producers are headquartered in the source country and it makes the host country a big labor reservoir that is inactive in the absence of offshoring.[10] Firms perform different tasks, which are combined in a Cobb-Douglas production technology to produce output y:

$$y(\omega) = \frac{z(\omega)}{1-z(\omega)} \exp\left[\frac{1}{z(\omega)}\int_0^{z(\omega)} \ln y(\omega, i) di\right], \tag{2}$$

where $y(\omega, i)$ denotes the output of task i and $z(\omega)$ is the length of the task interval (or the mass of tasks) performed by firm ω. The technology in Eq. (2) captures in a simple way the

[10] Admittedly, this is a fairly strong assumption. However, it is useful for our purposes, because it ensures without any further assumption that the host country is the low-wage country in our setting.

gains from labor division, as performing more tasks increases a firm's productivity. Assuming that task output equals labor input, the firm's total variable production costs are given by $C^v(\omega) = \int_0^{z(\omega)} \zeta(i)y(\omega,i)di$, where $\zeta(i)$ is the effective labor cost of task i, which is equal to the domestic wage w if a task is performed at home and equal to the foreign wage w^* multiplied by an iceberg trade cost parameter $\tau > 1$ if the task is performed abroad. Due to the underlying Cobb-Douglas technology, cost minimization establishes the result that expenditures are the same for all tasks. The marginal production cost of firm ω is then given by

$$c(\omega) = \begin{cases} [1 - z(\omega)]w & \text{if all tasks are produced at home} \\ [1 - z(\omega)]w\kappa^{s(\omega)} & \text{if share } s(\omega) \text{ of tasks is produced offshore} \end{cases}, \tag{3}$$

where $\kappa \equiv \tau w^*/w$ denotes the *effective* host country labor cost relative to the source country of offshoring. As we explain later, offshoring has fixed costs, and hence $\kappa < 1$ must hold in order to make it attractive for firms to shift task production abroad. Accordingly, we can associate the inverse of $\kappa^{s(\omega)}$ with the marginal *cost saving effect* of offshoring.

To enter the source country, firms must make an initial investment of $f_e > 0$ units of labor. This investment gives them a single draw from a task lottery and is immediately sunk. The outcome of the lottery is a technology tupel (z, s) with z and s being the length of the task interval and the share of tasks that are offshorable, respectively.[11] We consider a static model and, following Ghironi and Melitz (2005), abstract from fixed costs of production, so that all firms participating in the technology lottery start production, irrespective of their z-draw.[12] For tractability reasons, we assume that z is Pareto distributed over the unit interval with a probability density function $g_z(z) = k(1-z)^{k-1}$. The distribution of s is more sophisticated and potentially depends on a firm's z-level. We assume that a firm's probability to have at least some offshorable tasks is a positive function of the length of its task interval, and in the interest of tractability we set this probability equal to $\Pr_{z(\omega)}(s > 0) = \nu_0 + \nu_1 z$, with $\nu_0, \nu_1 \geq 0$ and $\nu \equiv \nu_0 + \nu_1 \in (0,1)$. Furthermore, for firms with some offshorable tasks, the share of tasks that can be put offshore, s, is uniformly distributed over the interval $(0, 1]$. Hence, for a firm with

[11] Alternatively, we could choose the unit cost of offshoring, τ, instead of the share of offshorable tasks, s, as the second variable contributing to firm heterogeneity. We decided against this alternative, because we use changes in the unit offshoring cost parameter in a comparative static analysis to shed light on how firms respond in their offshoring decision to a symmetric cost shock.

[12] We do not allow for selection into production, because our dataset contains many small producers, which employ only few domestic workers.

task length z, the ex ante expected value of s is given by $\mathbb{E}_{z(\omega)}[s] = (1/2)\text{Pr}_{z(\omega)}(s > 0)$.

Although the functional forms for modeling the distribution of the two random variables are admittedly somewhat restrictive, the chosen specifications are more flexible than they may appear at a first glance. In particular, they allow us to account for two observations from Figure 1. On the one hand, there is clustering of the data around the 'no offshoring event' and thus we observe a discrete share of non-offshoring firms in all deciles of the revenue distribution. This can be captured by $\nu_0 > 0$. On the other hand, there is a strong positive correlation between the probability to offshore and a firm's rank in the revenue distribution, which suggests that the distributions of the two productivity parameters may be not independent. This can be captured by $\nu_1 > 0$. Choosing a positive value of ν_1 is also akin to a simple probabilistic idea to offshoring. Relying on observations from Blinder and Krueger (2013) and Becker et al. (2013) that only a (relatively small) fraction of tasks can be classified as routine and do not require face-to-face communication – two prerequisites to count a task as offshorable – the probability of using at least some offshorable tasks is higher ceteris paribus for firms that use more tasks in their production process. In the empirical section, we show that positive values of ν_0 and ν_1 are supported by the data.

As an important feature of our model, we add the assumption that after the lottery firms are only informed about their z-level, but do not know how many (and which) of their tasks are offshorable. Under the veil of uncertainty they form expectations on the share of offshorable tasks s, conditional on probability $\text{Pr}_{z(\omega)}(s > 0)$, which is public knowledge. Depending on its expected s-level, $\mathbb{E}_{z(\omega)}[s]$, a firm with task length $z(\omega)$ can then choose to invest $f > 0$ units of labor into a fixed offshoring service input that provides information on the share s of offshorable tasks (as well as the type of tasks whose production can be conducted abroad). Being risk neutral, firms will make this offshoring investment only if its expected return is sufficiently high, and since the expected return is higher ceteris paribus for firms that have drawn a larger value of z in the lottery, our model establishes for a sufficiently high fixed cost parameter f selection of high-productive firms into offshoring. In Section 2.2, we identify the marginal offshoring firm that is characterized by task length $\hat{z}$ and is indifferent between making and not making the f investment, provided that there is selection of the best firms into offshoring. In Section 2.3, we will then determine the parameter domain that supports selection in our model.

2.2 Production costs and characterization of the marginal offfshoring firm

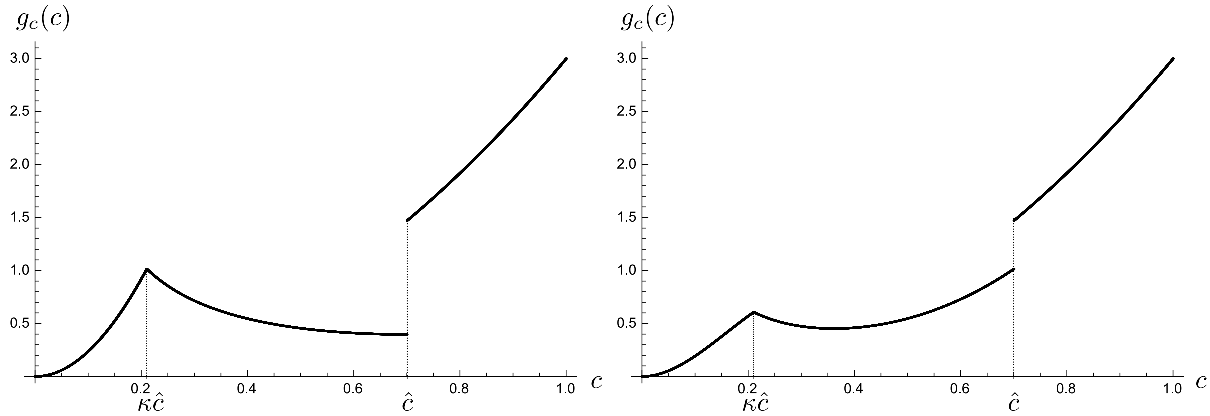
Even though our model features two forms of firm heterogeneity, we can conclude from Eq. (3) that their combined effect on firm-level performance measures is captured by a single variable: the marginal cost of production. This implies that we can learn about the distribution of firms in their various performance measures, when we understand how the distribution of the two technology parameter z and s maps into the distribution of marginal costs c. The marginal cost of non-offshoring firms is given by $c(\omega) = [1-z(\omega)]w$, according to Eq. (3). Non-offshoring firms are either low-productivity producers with task length $z(\omega) \leq \hat{z}$ or they are high productivity producers with task length $z(\omega) \geq \hat{z}$ and no offshorable task. Due to the inverse link between $c(\omega)$ and $z(\omega)$, there is no difference between ranking non-offshoring firms by their task length or the marginal costs – with the ordering of firms flipped – and we can directly infer the distribution of marginal costs c from the distribution of task length z.

Things are more complicated for offshoring firms, which are high-productivity firms with task length of $z(\omega) \geq \hat{z}$ whose production process includes at least some offshorable tasks. The marginal cost of an offshoring firm is given by $c(\omega) = [1 - z(\omega)]w\kappa^{s(\omega)}$, according to Eq. (3) and thus the product of two random variables. Therefore, the ranking of $c(\omega)$ cannot be inferred from the ranking of $z(\omega)$. Characterizing the distribution of marginal cost in the population of offshoring firms becomes even more sophisticated, if the probability of offshoring $\Pr_z(s > 0)$ depends on the number of tasks operated by the firm, i.e. if $\nu_1 > 0$, because in this case the distributions of z and s are not independent. In the appendix, we show how we can link the distributions of z and s to compute the probability density function (pdf) of marginal production cost c:

$$g_c(c) = \begin{cases} (1-\nu+\nu_1 c)kc^{k-1} - \frac{1}{\ln\kappa}\left\{\nu c^{k-1}\left[\left(\frac{1}{\kappa}\right)^k - 1\right] - \nu_1 \frac{kc^k}{k+1}\left[\left(\frac{1}{\kappa}\right)^{k+1} - 1\right]\right\} & \text{if} \quad c \leq \kappa\hat{c} \\ (1-\nu+\nu_1 c)kc^{k-1} - \frac{1}{\ln\kappa}\left\{\nu c^{k-1}\left[\left(\frac{\hat{c}}{c}\right)^k - 1\right] - \nu_1 \frac{kc^k}{k+1}\left[\left(\frac{\hat{c}}{c}\right)^{k+1} - 1\right]\right\} & \text{if} \quad c \in (\kappa\hat{c}, \hat{c}] \\ kc^{k-1} & \text{if} \quad c > \hat{c} \end{cases}, \tag{4}$$

with $\hat{c} \equiv (1 - \hat{z})w$. The probability density function of c is illustrated for two different sets of parameters in Figure 2.

As we can see from Eq. (4) and Figure 2 the pdf of marginal costs, $g_c(c)$, has support on



Parameter values: $k = 3$, $\hat{c} = 0.7$, $\kappa = 0.3$, and $\nu_0 = 0.7, \nu_1 = 0.1$ (left panel); $\nu_0 = 0.1, \nu_1 = 0.7$ (right panel).

Figure 2: The probability density function $g_c(c)$

the unit interval and features a discontinuity at $\hat{c}$. This is because for firms with task length $z(\omega) \geq \hat{z}$ investment into offshoring is attractive, and a subset of these firms detects to use at least some offshorable tasks and thus starts offshoring. Since offshoring firms experience a marginal cost saving and are thus shifted to a lower $c(\omega)$ and since the fraction of firms that is affected by this cost saving is discrete for any $z > 0$ ($\nu = \nu_0 + \nu_1 > 0$), selection into offshoring generates a discontinuity of the pdf at $\hat{c} < 1$ in Figure 2. The kink of the pdf function at $\kappa\hat{c}$ is also rooted in the selection of high-productivity firms into offshoring. More specifically, since firms with $z(\omega) < \hat{z}$ refuse to make the fixed cost investment for learning about the offshorability of their tasks, none of these firms is shifted towards lower marginal costs. This imposes a binding (selection) constraint on the number of firms that can be located at the marginal cost interval $[\kappa\hat{c}, \hat{c}]$. For marginal costs $c < \kappa\hat{c}$ the selection constraint is not binding, because the maximum possible cost saving from offshoring when shifting all tasks abroad is given by κ, and hence a firm with task length $z(\omega) < \hat{z}$ could not be shifted to a marginal cost lower than $\kappa\hat{c}$ even if it would make the investment into offshoring, despite an expected profit loss.

We can now make use of the well established result that in a model like ours firms maximize their profits by setting prices as a constant markup $\sigma/(\sigma - 1)$ over their marginal costs, $c(\omega)$: $p(\omega) = c(\omega)\sigma/(\sigma - 1)$. Since firm-level revenues are given by $r(\omega) = I[p(\omega)/P]^{1-\sigma}$, we can express relative revenues of two firms as a decreasing function of their marginal cost differential:

$$\frac{r(\omega_1)}{r(\omega_2)} = \left[\frac{c(\omega_1)}{c(\omega_2)}\right]^{1-\sigma}. \tag{5}$$

Acknowledging this monotonic relationship, the distribution of revenues can be described by the pdf in Eq. (4). The link between revenues and marginal costs in Eq. (5) describes an ex post relationship, i.e. the relationship between marginal costs and revenues after the investment decision of firms has been made.

The *relative* revenue increase a firm with share $s(\omega)$ of offshorable tasks can realize under offshoring is given by $\kappa^{s(\omega)(1-\sigma)}$. To achieve this revenue increase, firms must invest the fixed amount of labor, $f > 0$, under the veil of uncertainty, and therefore form rational expectations regarding their gains from offshoring prior to this investment. The expected relative revenue gain from offshoring depends on the cost saving under all possible realizations of s, and for $\nu_1 > 0$ it is larger for firms with a better z-draw: $\mathbb{E}_{z(\omega)}\left[\kappa^{s(1-\sigma)}\right] = \Pr_{z(\omega)}(s > 0)\int_0^1 \kappa^{s(1-\sigma)}ds$, with $d\text{PR}_z(s > 0)/dz > 0$. In *absolute* terms, there is a second advantage that renders offshoring more attractive for firms with a better z-draw. They make higher revenues at any possible realization of s. This can be seen when looking at the revenue differential of two non-offshoring firms: $\{[1 - z(\omega_1)]/[1 - z(\omega_2)]\}^{1-\sigma}$. Accordingly, firms with a larger z can in expectation more easily cover the fixed cost of offshoring, fw. Acknowledging constant markup pricing, the expected profit gain from offshoring of a firm with task length z can then be expressed as $\Pr_{z(\omega)}(s > 0)[1 - z(\omega)]^{1-\sigma}r(1)[\int_0^1 \kappa^{s(1-\sigma)}ds - 1]/\sigma - fw$, where $r(1)$ is the revenue of the least productive firm with $z = 0$ and $c = w$, which is a firm that does not offshore. The marginal offshoring firm with task length $\hat{z}$ or marginal cost $\hat{c} = (1 - \hat{z})w$ is indifferent between making and not making the investment of fw and is therefore characterized by the following condition

$$\sigma f w = \left(\nu - \nu_1 \frac{\hat{c}}{w}\right)\left(\frac{\hat{c}}{w}\right)^{1-\sigma} r(1)\left[\frac{\kappa^{1-\sigma} - 1}{(1-\sigma)\ln\kappa} - 1\right]. \tag{6}$$

Since Eq. (6) summarizes the solution to the investment problem of firms and since ex post relative firm-level variables are characterized by relative marginal costs of firms, we can omit firm index ω from now on and index firms by their marginal cost c.

2.3 The general equilibrium

To solve for the general equilibrium, we choose source country labor as numéraire and set the source country wage rate equal to one: $w = 1$. Using Eq. (4), we can then compute economy-wide revenues according to

$$R = M\int_0^1 r(c)g_c(c)dc$$
$$= Mr(1)\left[\frac{k}{k-\sigma+1} + \hat{c}^{k-\sigma+1}\left(\frac{k\nu}{k-\sigma+1} - \frac{k\nu_1\hat{c}}{k-\sigma+2}\right)\left(\frac{\kappa^{1-\sigma}-1}{(1-\sigma)\ln\kappa} - 1\right)\right], \tag{7}$$

where $k > \sigma - 1$ has been assumed to ensure a finite positive value of R (see the appendix for derivation details). Since free entry implies that in expectation firms must make zero profits (net of lottery fixed costs), the economy-wide return to production $R/\sigma - M\hat{c}^k f$ must equal total fixed market entry costs Mf_e. Total revenues are therefore proportional to total (market entry plus offshoring) fixed cost expenditures, $R = M\sigma\left(f_e + \hat{c}^k f\right)$. Together with Eqs. (6) and (7), this establishes a relationship between the marginal cost of the offshoring firm that is indifferent between making and not making investment f, $\hat{c}$, and the effective wage differential between the host and the source country of offshoring, κ, which we call *offshoring indifference condition* (OC):

$$\Gamma_1(\hat{c},\kappa) = \frac{\hat{c}^{\sigma-1}}{\nu-\nu_1\hat{c}}\frac{k}{k-\sigma+1} + \left\{\frac{\hat{c}^k}{\nu-\nu_1\hat{c}}\left[\frac{(\sigma-1)\nu}{k-\sigma+1} - \frac{(\sigma-2)\nu_1\hat{c}}{k-\sigma+2}\right] - \frac{f_e}{f}\right\}\left[\frac{\kappa^{1-\sigma}-1}{(1-\sigma)\ln\kappa} - 1\right] = 0.$$

As formally shown in the appendix, $\Gamma_1(\cdot) = 0$ establishes a negative link between $\hat{c}$ and κ. The larger the effective labor costs in the host country relative to the source country of offshoring are, the smaller is the cost saving effect of offshoring and the more productive the marginal firm that is indifferent between investing and not investing f must be in order to avoid (in expectation) losses from offshoring. Intuitively, if the cost saving from offshoring vanishes due to $\kappa = 1$, all firms prefer domestic production, resulting in $\hat{c} = 0$. In contrast $\hat{c}$, reaches a maximum at low levels of κ.

A second link between $\hat{c}$ and κ can be determined, when noting from above that free entry into the technology lottery implies that all income accrues to workers. Since global income is equal to total consumption expenditures, we have $R = L + w^*L^*$. Furthermore, constant markup pricing establishes the well-known result that variable production costs are a constant fraction

$(\sigma-1)/\sigma$ of a firm's revenues, with part of these costs accruing to imported tasks. As formally shown in the appendix, the wage bill for workers in the host country of offshoring can thus be expressed as a function of aggregate revenues according to

$$w^*L^* = R\frac{\sigma-1}{\sigma}\frac{\hat{c}^{k-\sigma+1}\left(\nu-\nu_1\hat{c}\frac{k-\sigma+1}{k-\sigma+2}\right)\frac{1+\kappa^{1-\sigma}[(1-\sigma)\ln\kappa-1]}{[(1-\sigma)\ln\kappa]^2}}{1+\hat{c}^{k-\sigma+1}\left(\nu-\nu_1\hat{c}\frac{k-\sigma+1}{k-\sigma+2}\right)\left(\frac{\kappa^{1-\sigma}-1}{(1-\sigma)\ln\kappa}-1\right)}. \tag{8}$$

In combination with $R = L + w^*L^*$ this establishes a second implicit link between the two endogenous variables κ and $\hat{c}$, which reflects adjustments in the effective wage differential in response to changes the attractiveness of offshoring that are enforced by labor market clearing in the two economies:

$$\Gamma_2(\kappa,\hat{c}) \equiv \kappa\left\{\frac{\sigma}{\sigma-1}\frac{1+\hat{c}^{k-\sigma+1}\left(\nu-\nu_1\hat{c}\frac{k-\sigma+1}{k-\sigma+2}\right)\left(\frac{\kappa^{1-\sigma}-1}{(1-\sigma)\ln\kappa}-1\right)}{\hat{c}^{k-\sigma+1}\left(\nu-\nu_1\hat{c}\frac{k-\sigma+1}{k-\sigma+2}\right)\frac{1+\kappa^{1-\sigma}[(1-\sigma)\ln\kappa-1]}{[(1-\sigma)\ln\kappa]^2}}-1\right\}-\frac{\tau L}{L^*}=0 \tag{9}$$

We refer to this implicit relationship by the term *labor market constraint* (LC) and formally show in the appendix that $\Gamma_2(\cdot)=0$ establishes a positive link between κ and $\hat{c}$. The larger is $\hat{c}$, the more firms are engaged in offshoring and the larger is ceteris paribus the demand for foreign workers. This drives up foreign wages and increases κ. If $\hat{c}$ falls to zero, there is no offshoring and, lacking access to occupations outside the production sector, wages in the host country and thus also κ fall to zero. In contrast, κ reaches a maximum at a high level of $\hat{c}$.

The equilibrium values of $\hat{c}$ and κ are jointly determined by the offshoring indifference condition and the labor market constraint. Thereby, our model features a unique interior equilibrium if offshoring cost parameters τ and f are sufficiently high.[13] The impact of changes in the two offshoring cost parameters is illustrated in Figure 3, which depicts the two loci OC and LC for different parameter configurations. A higher variable offshoring cost parameter τ implies for a given volume of offshoring that more foreign workers must be employed in order to provide the required amount of tasks for production in the source country. Therefore, the effective cost for employing foreign relative to domestic labor, κ, must increase to restore labor market clearing. This effect is captured by a counter clockwise rotation of locus LC in Figure 3, which makes an interior solution with intersection of OC and LC at $\hat{c} < 1$ and $\kappa < 1$ more likely. A higher

[13]The critical levels of $\tau L/L^*$ and f depend – among other model parameters – on the levels of ν_0 and ν_1. In the special case of $\nu_0 = 0$, a unique interior equilibrium is guaranteed for any combination of $\tau L/L^*$ and f.

offshoring fixed cost parameter makes offshoring less attractive ceteris paribus and therefore lowers the cutoff cost level characterizing the firm that is indifferent between making and not making the investment of f. This effect is captured by a clockwise rotation of locus OC in Figure 3, which also makes the existence of an interior equilibrium more likely.

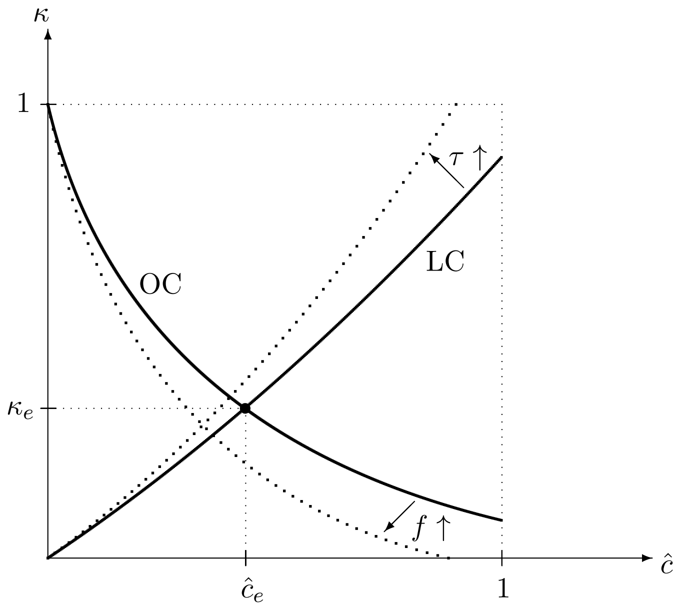


Figure 3: Equilibrium values of $\hat{c}$ and κ

In an interior equilibrium as captured, for instance, by the intersection point of the solid OC and LC loci, an increase in either offshoring cost parameter lowers the cutoff cost level $\hat{c}$ and thus the share of offshoring firms in our model. The consequences of higher offshoring costs on the effective wage differential κ depend on which offshoring cost parameter changes. If the fixed offshoring cost parameter increases, the provoked fall in host country labor demand unambiguously lowers the effective wage differential κ. Whereas this labor demand effect is also present when the variable offshoring parameter increases, it is counteracted and dominated by the initial increase in τ, so that the effective wage differential increases.

2.4 Offshoring margins and welfare

With the general equilibrium outcome at hand, we can look in more detail at the adjustments of offshoring along two margins that play a prominent role in the recent trade literature: the extensive margin, capturing changes in the mass of offshoring firms; and the intensive margin, capturing changes in the volume of offshoring by incumbent offshoring firms. Looking at the extensive margin first, we can note that the share of firms that can offshore is c-specific and depends on the firm's endogenous decision on whether to make investment f or not. Denoting the share of offshoring firms relative to all firms with the same marginal cost c by $\chi(c)$, we can compute

$$\chi(c) = \begin{cases} 1 - \left[1 - \frac{1}{\ln \kappa} \frac{\nu\left[\left(\frac{1}{\kappa}\right)^k - 1\right] - \nu_1 \frac{kc}{k+1}\left[\left(\frac{1}{\kappa}\right)^{k+1} - 1\right]}{(1-\nu+\nu_1 c)k}\right]^{-1} & \text{if} \quad c \leq \kappa\hat{c} \\ 1 - \left[1 - \frac{1}{\ln \kappa} \frac{\nu\left[\left(\frac{\hat{c}}{c}\right)^k - 1\right] - \nu_1 \frac{kc}{k+1}\left[\left(\frac{\hat{c}}{c}\right)^{k+1} - 1\right]}{(1-\nu+\nu_1 c)k}\right]^{-1} & \text{if} \quad c \in (\kappa\hat{c}, \hat{c}] \\ 0 \end{cases} \tag{10}$$

according to Eq. (4). It is easily confirmed that $\chi'(c) < 0$ holds for all $c < \hat{c}$, implying that the share of offshoring firms decreases in c. The economy-wide share of offshoring firms can be computed according to

$$\chi = \int_0^1 \chi(c)g(c)dc = \hat{c} - \int_0^{\hat{c}} (1 - \nu + \nu_1 c)kc^{k-1}dc = \hat{c}^k \left[\nu - \nu_1 \frac{k}{k+1}\hat{c}\right] \tag{11}$$

and is increasing in the cutoff level of marginal costs $\hat{c}$: $d\chi/d\hat{c} = k\hat{c}^{k-1}(\nu - \nu_1\hat{c}) > 0$. Since we know from Figure 3 that $d\hat{c}/df < 0$ and $d\hat{c}/d\tau < 0$, we can thus conclude that a decline in either offshoring cost parameter increases the share of offshoring firms and thus raises offshoring along the extensive margin.

To study adjustments of offshoring along the intensive margin, we can note that total task expenditures of offshoring firms are given by $[(\sigma - 1)/\sigma]R[1 - R^d/R]$, with

$$\frac{R^d}{R} = \frac{1 - \hat{c}^{k-\sigma+1}\left(\nu - \nu_1\hat{c}\frac{k-\sigma+1}{k-\sigma+2}\right)}{1 + \hat{c}^{k-\sigma+1}\left(\nu - \nu_1\hat{c}\frac{k-\sigma+1}{k-\sigma+2}\right)\left(\frac{\kappa^{1-\sigma}-1}{(1-\sigma)\ln\kappa} - 1\right)}, \tag{12}$$

being the fraction of aggregate revenues accruing to domestic producers. In view of Eq. (8) we can thus write the economy-wide expenditure share for imported tasks as follows

$$\rho \equiv \frac{w^*L^*}{[(\sigma-1)/\sigma]R[1 - R^d/R]} = \frac{1 + \kappa^{1-\sigma}[(1-\sigma)\ln\kappa - 1]}{[(1-\sigma)\ln\kappa](\kappa^{1-\sigma} - 1)} = \frac{\kappa^{1-\sigma}}{\kappa^{1-\sigma} - 1} - \frac{1}{(1-\sigma)\ln\kappa}, \tag{13}$$

with $\lim_{\kappa\to 0}\rho = 1$, $\lim_{\kappa\to 1}\rho = 1/2$, and $d\rho/d\kappa < 0$. From this we can conclude that incumbent offshoring firms expand their expenditure share for imported tasks if the effective cost of employing foreign labor, κ, decreases. From Figure 3 we can infer that $d\kappa/d\tau > 0$ and $d\kappa/df < 0$, and hence the response of offshoring to exogenous changes in the offshoring cost parameters along the intensive margin depends on the specific nature of the cost change. If the variable cost of offshoring decreases, the effective cost of foreign labor decreases despite an increase in the foreign wage rate, and this triggers an expansion of offshoring along the intensive margin which complements the increase in offshoring along the extensive margin. If, however, the fixed cost of offshoring decreases, the effective cost of foreign labor increases with host country wages so that the increase of offshoring along the extensive margin is counteracted by a decline of offshoring along the intensive margin.

A distinction between the extensive and intensive margin of offshoring is important for gaining a better understanding about the welfare implications of offshoring in the source country.[14] Since preferences are homothetic, we can use the representative consumer in a normative interpretation and consider per-capita labor income ($\widehat{=}$ GDP per capita) as our preferred welfare measure. In view of $w = 1$, we can thus express source country welfare as the inverse of the consumer price index: $W = P^{-1}$. To determine the consumer price index, we can start from the observation that revenues of the least productive (non-offshoring) firm, which faces a marginal cost equal to one, is given by $r(1) = p(1)y(1)$. Accounting for Eq. (1) and our previous insight that global consumption expenditure is equal to total source and host country labor income $L + w^*L^*$, revenues of the least productive firm can be expressed as

$$r(1) = \frac{L + w^*L^*}{P^{1-\sigma}} \left(\frac{\sigma - 1}{\sigma}\right)^{\sigma-1}. \tag{14}$$

A second expression for the revenues of the least productive producer can be found when combining the indifference condition of the marginal offshoring firm in Eq. (6) with the offshoring indifference condition $\Gamma_1(\cdot) = 0$:

$$r(1) = \sigma f \left[\frac{f_e}{f} - \frac{\hat{c}^k}{\nu - \nu_1\hat{c}} \left(\nu\frac{\sigma - 1}{k - \sigma + 1} - \nu_1\hat{c}\frac{\sigma - 2}{k - \sigma + 2}\right)\right] \frac{k - \sigma + 1}{k}. \tag{15}$$

[14] Modeling the host country in a parsimonious way, our model lacks important features that makes it a suitable vehicle for studying host country welfare. Therefore, we focus on the source country of offshoring in the welfare analysis.

The two Eqs. (14) and (15) jointly determine price index P and thus source country welfare

$$W = \left\{ \frac{L + \kappa L^*/\tau}{\sigma f} \left[\frac{f_e}{f} - \frac{\hat{c}^k}{\nu - \nu_1 \hat{c}} \left(\nu \frac{\sigma - 1}{k - \sigma + 1} - \nu_1 \hat{c} \frac{\sigma - 2}{k - \sigma + 2} \right) \right]^{-1} \frac{k}{k - \sigma + 1} \right\}^{\frac{1}{\sigma - 1}} \frac{\sigma - 1}{\sigma}. \quad (16)$$

A decline in τ induces an expansion of offshoring along both the intensive and extensive margin and therefore raises foreign labor demand. Whereas this leads to higher foreign wages, the increase in the foreign wage rate is not strong enough to dominate the initial decline in the variable offshoring cost. As a consequence, the effective foreign labor cost decreases, reflecting an appreciation of domestic relative to foreign labor and thus an improvement in the source country's (double) factorial terms of trade, with positive welfare consequences. Things are different if the fixed cost of offshoring decreases, because the expansion of offshoring along the extensive margin not only raises foreign wages but also the relative effective cost of employing workers in the host country. This induces a decline of offshoring along the intensive margin and worsens the (double) factorial terms of trade of the source country. The depreciation of domestic relative to foreign labor may be strong enough to dominate the source country's direct welfare gain from a lower offshoring fixed cost. In the appendix, we provide a formal discussion of these effects and illustrate the possibility of welfare losses from a lower fixed offshoring cost by means of a numerical example.

Welfare in the source country and the relative importance of the extensive and the intensive margin of offshoring are the two main targets of the empirical analysis conducted below. There, we use the formal structure of our model as guidance to estimate the main parameters of this model. With the parameter estimates at hand, we will then analyze the aptitude of our model to capture important data features, quantify the welfare effects of offshoring, and shed light on how important accounting for the overlap actually is for capturing the welfare effects of offshoring and the relative importance of the two margins of offshoring.

3 A quantitative exercise

To make our model accessible to a quantitative analysis, we collect data from two different sources and estimate key parameters of our model using a structural approach. In the next two subsections, we describe the data and outline the empirical methodology. The parameter estimates from our quantitative exercise are reported and discussed in Subsection 3.3, and in Subsection 3.4 we use these estimates to quantify the welfare effects of offshoring for Germany.

3.1 Data sources and descriptives

We use data on revenues and offshoring of German establishments from the Establishment Panel of the Institute of Employment Research (IAB) in Nuremberg. This database reports detailed establishment information from employer surveys at an annual basis since 1993.[15] However, information on the offshoring activity of German establishments is only available in the years 1999, 2001, and 2003. Following Moser et al. (2015), we associate offshoring with the purchase of intermediates or other inputs from abroad in the previous business year. Dropping establishments that lack either offshoring or revenue information, gives us a sample of 12,250 different establishments and 20,341 establishment observations for the three years. We do not exploit the time-series variation in the data and rely on the observational information to construct a cross section of German establishments for estimating the parameters of the static model outlined in Section 2.[16]

Building on the idea of task offshoring it is important for our structural approach to gather information on the number of tasks performed in German establishments. To construct this data, we rely on the BIBB/BAuA Employment Survey, which provides information on workplace characteristics, including the tasks performed and the occupations held by respondents, for a representative sample of German employees with a working time of more than 10 hours per week (see Rohrbach-Schmidt, 2009, for a detailed description). Interviews have been conducted six times at irregular intervals since 1979. Given the temporal proximity to the offshoring events in the IAB Establishment Panel, we use information on tasks and occupations from the 2006 survey, which covers 20,000 employees. In this survey, we can distinguish 28 different tasks (activities), which are listed in the data appendix. Depending on the content, interviewees can answer the 28 questions on whether they perform a certain task either with *often/sometimes/never* or with *yes/no*. We map a task to an occupation if at least 60 percent of the respondents in an occupation

[15] Although the sample of establishments included in the survey itself is disproportionately stratified, the sample can be made representative for the total population of German establishments, when using the appropriate weighting schemes offered by the Research Data Centre at the IAB (cf. Fischer et al., 2008).

[16] The IAB Establishment Panel provides plant-level information, which unfortunately cannot be aggregated to firm-level information with the available data. To acknowledge that the observational unit in the empirical exercise are plants, we use the term establishments instead of firms from now on.

declared to perform that task to some extent – i.e. if their answer was *often*, *sometimes*, or *yes*. We distinguish 341 different occupation units (*Berufsordnungen*), using the classification of the Federal Employment Agency, KldB 1988, and collect for each of these occupations the number of tasks performed, according to the responses of the interviewees.[17]

To aggregate the thus constructed task content of occupations to the establishment level, we use the Linked Employer-Employee (LIAB) database of the Institute of Employment Research, which provides record linkages for matching detailed administrative data on employees registered with the German social security system – including information on their occupations – to the IAB Establishment Panel. This matching procedure allows us to extract knowledge about the task content of production of German establishments from the occupations of their workforce. Following this procedure we can match 302 out of the 341 occupations to the IAB Establishment Panel and find only 5 establishments or 7 establishment observations performing zero tasks. Whereas an outcome with zero tasks is consistent with our model, the cases are suspicious, because all of the establishments with zero tasks have only one type of occupation: laundry workers and pressers. Since activities in this occupation are not captured by the 2006 BIBB-BAuA Employment Survey, we have decided to drop these establishments from our dataset, and hence end up with a total number of 12,245 establishments and 20,334 establishment observations. For consistency with our theoretical model, we need a measure of the length of the task interval ranging between 0 and 1 and obtain this by dividing the number of tasks in each firm by the total number of tasks reported by the BIBB/BAuA survey. Table 1 reports the descriptives of establishments in our dataset.

From the first row of Table 1, we see that 22.7 percent of the establishments in our dataset offshore. This figure is higher than the share of offshoring firms reported by Moser et al. (2015). The main reason for this difference is that Moser et al. (2015) use the time-variation in the IAB Establishment Panel and define offshoring by a (qualitative) increase of an establishment's share of foreign intermediates in two consecutive periods with offshoring information. This is a more restrictive definition of offshoring than the one we use in our cross section. The second

[17]By using information from the BIBB/BAuA Employment Survey to learn about the task content of occupations, we follow Spitz-Oener (2006) and Becker et al. (2013). In a recent paper, Becker and Muendler (2015) have used this database to shed light on how globalization has changed the task composition of occupations over the last decades. It is notable that Becker and Muendler (2015) use slightly different task definitions and distinguish only 15 activities which they can observe in different periods. As shown by a robustness check in the appendix, using these alternative task definitions would not change our main results.

Table 1: Descriptive Statistics

	Mean	Median	Std. Dev.	Min.	Max.
Offshoring	0.227	0.000	0.419	0.000	1.000
Task interval	0.298	0.286	0.142	0.036	0.893
Tasks per occupation	5.241	5.231	1.665	0.192	16.000
Workforce with soc. sec.	12.257	3.000	88.833	0.000	42,291
Revenues in million Euro	2.461	0.347	35.556	0.004	12,055

Descriptives are computed based on 20,334 establishment observations over the years 1999, 2001, 2003, using inverse probability weights to make the sample of establishments representative, as suggested by the Research Data Centre at the IAB.

row tells us that the establishments in our dataset are nicely distributed over the unit task interval with positive skew. In the third row we see that the typical occupation in a German establishment is a multi-task entity and that the employees perform more than five tasks in an average establishment. The finding that some establishments have less than one task per occupation indicates that there are a few occupations, which we were not able to link to at least one task and which accordingly have a task content of zero. Since we consider all establishments, for which we have the required offshoring, revenue, and task information, our dataset features large differences of establishments in both the size of workforce and the size of revenues.[18]

3.2 Estimation strategy

To estimate the parameters of our model, we implement a minimum distance methods-of-moments (MM) approach as outlined in Ferguson (1958) and Cameron and Trivedi (2005). This approach is similar to other MM applications and builds on the idea to specify a vector of n_m observed population moments, $\mathbf{m}$, which is linked to a vector of n_x structural parameters of the model, $\mathbf{x}$, according to $\mathbf{m} = \mu(\mathbf{x})$, where $\mu(\mathbf{x})$ is a $n_m \times 1$ vector function. If the number of moments, n_m, is (weakly) larger than the number of structural parameters, n_x, we can estimate the structural parameters $\mathbf{x}$ by minimizing the weighted (squared) distance between observed moments $\mathbf{m}$ and computed moments $\mu(\mathbf{x})$, subject to a vector of constraints, **Cons** that are

[18] Since the aggregation of occupations to establishments using workforce information from LIAB is also possible if the employment status is not subject to social security payments, our dataset includes small establishments that do not have a single employee for whom they have a legal obligation to make such payments.

imposed by the theoretical model:

$$\widehat{\mathbf{x}}_{MD} = \text{argmin}_{\mathbf{x}} \ (\widehat{\mathbf{m}} - \mu(\mathbf{x}))' \mathbf{W} (\widehat{\mathbf{m}} - \mu(\mathbf{x})), \quad s.t. \quad \mathbf{Cons}, \tag{17}$$

where $\mathbf{W}$ is a $n_m \times n_m$ positive-semidefinite weighting matrix and a hat indicates observed or estimated variables. The specific assumption of the MM estimator considered here is that $\mathbf{m}$ is a vector of reduced form parameters, whose estimates $\widehat{\mathbf{m}}$ are the means of subsets of observations. The close relationship between the MM approach of Ferguson (1958) and GMM is obvious from inspection of Eq. (17). As pointed out by Hall (2005), "the statistical framework developed by Ferguson (1958) contains many of the elements which reappeared in the GMM literature twenty-five years later" (p. 11). The main difference is that Ferguson's (1958) approach is more restrictive. To see this, we can specify a general vector function $\mathbf{\Delta}(\mathbf{o}, \mathbf{x})$ on observables $\mathbf{o}$ and structural parameters $\mathbf{x}$, which captures the data generating process. GMM requires that the moment conditions fulfill $\mathbb{E}[\mathbf{\Delta}(\mathbf{o}, \mathbf{x})] = 0$. In our MM estimation, we presume the functional form $\Delta_t(\mathbf{o}, \mathbf{x}) \equiv m_t(\mathbf{o}) - \mu_t(\mathbf{x})$, for all moment conditions $t = \{1, ..., n_m\}$, where $m_t(\mathbf{o})$ is the mean of a subset of observations $\mathbf{o}$. Accordingly, the MM approach considered here can be interpreted as GMM with a specific functional relationship between the observables and structural parameters that conditions on reduced form parameters $\mathbf{m}$.

A natural candidate for weighting the moment conditions in Eq. (17) is the identity matrix, which puts equal weights on all these conditions. However, this is not recommendable from an efficiency point of view, because this simple weighting matrix does not account for differences in the quality of measurement of reduced form parameters m_t. Therefore, the literature suggests to specify a diagonal weighting matrix relying on inverse variances of the reduced form parameter estimates. This approach puts higher weight on more precisely measured moments and is the optimal weighting matrix for given reduced form estimates $\widehat{\mathbf{m}}$ (cf. Cameron and Trivedi, 2005).[19]

[19] As a third alternative, one could rely on the variance-covariance matrix of the moment conditions $\mathbf{\Delta}(\mathbf{o}, \mathbf{x})$ to specify the weighting matrix, as suggested by Hansen (1982). This approach does not condition on reduced form estimates and is the preferred choice for GMM applications with more moment conditions than parameters. Since the variance-covariance matrix of moment conditions is unobservable, one has to estimate it. Estimation therefore requires at least two steps – the first one for determining the weighting matrix and the second one for estimating the parameters – which in our application is very time-consuming. From econometric theory we know that the choice of the weighting matrix affects efficiency but not consistency of the parameter estimates and that in the case of finite samples standard errors tend to be downward biased even when using the asymptotically optimal weighting matrix (cf. Windmeijer, 2004). Therefore, we prefer a parsimonious approach and use inverse variances of the reduced form parameter estimates to specify a diagonal weighting matrix. In a robustness analysis presented in Section 6 we analyze how our results change when relying on an estimate of the inverse variance-covariance matrix of moment conditions for constructing $\mathbf{W}$.

We apply the MM estimator to determine three parameters of our model, namely ν_0, ν_1, and σ, and we compute the corresponding values of k and κ using two constraints imposed by our model: the adding up condition for the share of offshoring firms in Eq. (11); and the adding up condition for the share of aggregate revenues of non-offshoring establishments relative to all producers in Eq. (12). We also impose $k, \sigma > 1$, $k > \sigma - 1$, and $\nu = \nu_0 + \nu_1 \leq 1$ as additional parameter constraints from our model. Since the moment conditions outlined below are highly nonlinear functions of the parameters of our model, we cannot solve the minimization problem in Eq. (17) analytically. Therefore, we choose a numerical approach considering a discrete parameter space with fine grid and compute the theory moments for each parameter cell.

To construct the parameter space, we pick in a first step different combinations of ν_0, ν_1 from the unit interval in step length 0.01. Due to constraint $\nu_0 + \nu_1 \leq 1$, this gives 5,148 possible (ν_0, ν_1)-pairs, for which we then determine the corresponding levels of k. To do so, we identify the marginal offshoring establishment, i.e. the offshoring producer with the smallest revenue in the data, and count all non-offshoring establishments with revenues lower than that. As a result of this exercise, we find that none of the smallest 18 establishments in our dataset offshores and the quantile position of the marginal offshoring establishment in the revenue distribution therefore corresponds to a value of 0.001. This confirms the insight from Figure 1 that offshoring is prevalent in all size categories of German establishments. Since our model produces an inverse relationship between the rank of firms in the revenue distribution and their rank in the marginal cost distribution, we can express the cutoff cost level as a function of k: $\hat{c} = 0.999^{1/k}$. We can then determine a unique value of χ for all 5,184 possible combinations of ν_0 and ν_1 using Eq. (11) and can equate the thus computed χ-value with the observed share of offshoring firms in our dataset, $\hat{\chi} = 0.227$. Thereby, we constrain the possible solutions for k to values larger than one: $k \geq 1+10^{-7}$. Depending on the specific combination of parameters ν_0 and ν_1, the solutions for parameter k can vary between a low value of 1.029 and a large value of 111.017.

In a second step, we select the possible values of σ in step length 0.01 that fulfill the parameter constraints $\sigma \geq 1+10^{-7}$ and $\sigma \leq k+1-10^{-7}$. For all possible combinations of the four parameter values ν_0, ν_1, σ, and k, we then determine the corresponding level of $\kappa \in (0, 1)$ by equating R^d/R from Eq. (12) with its observed counterpart $\widehat{R^d/R} = 0.430$. Although this problem does not have an interior solution for all possible parameter combinations, we can still fill 921,758 out of the 1,311,863 remaining cells of the parameter space. For these cells, we find a solution that fulfills all relevant parameter constraints, and we use the thus determined parameter values as input to construct the computed moments, which can then be combined with the reduced form parameter estimates to numerically solve the minimization problem in Eq. (17).

For the MM approach, the number of moments must be at least as high as the number of estimated parameters. We therefore choose four data moments, which we aim to target with our model.[20] As a first target, we consider the variance of marginal costs, which in turn are inversely related to the number of tasks performed in an establishment. Since marginal costs are not directly observable and information on the tasks performed in an occupation is only available for the German workforce, we restrict our analysis to non-offshoring establishments and compute for these producers the variance of marginal costs from information of the task content of production, according to $c = 1 - z$. The variance of marginal costs of non-offshoring firms gives the first data moment for our estimation, which we denote by m_1. Accounting for Eq. (4), we can compute the theoretical counterpart to this data moment as follows:

$$\mu_1(\nu_0, \nu_1, k) = \frac{k}{k+2} \frac{1 - \hat{c}^{k+2}\left(\nu - \nu_1 \frac{k+2}{k+3}\hat{c}\right)}{1 - \hat{c}^{k}\left(\nu - \nu_1 \frac{k}{k+1}\hat{c}\right)} - \left[\frac{k}{k+1} \frac{1 - \hat{c}^{k+1}\left(\nu - \nu_1 \frac{k+1}{k+2}\hat{c}\right)}{1 - \hat{c}^{k}\left(\nu - \nu_1 \frac{k}{k+1}\hat{c}\right)}\right]^2, \tag{18}$$

with $\nu = \nu_0 + \nu_1$ and $\hat{c} = 0.999^{1/k}$.

To construct the second moment, we make use of the insight from Eq. (5) that revenues of establishments with $c < 1$ are linked to the revenues of the least productive producer with marginal cost $c = 1$ by means of $\ln r(c) = \ln r(1) + (1 - \sigma)\ln(c)$. For non-offshoring firms we can infer the level of marginal costs from the observed usage of tasks, according to $c = 1 - z$. This allows us to compute average log marginal costs of non-offshoring firms: $\mathbb{E}(\ln \hat{c}|\text{not offshoring}) = -0.576$. Since the smallest establishments in our dataset have exceptionally low revenues, our estimates may be biased by outliers when simply relying on the lowest revenue in our dataset for measuring $r(1)$. To overcome this problem, we approximate the revenue of the marginal establishment by the revenue of the establishment at the first percentile. This is not problematic from a theory point of view as long as this establishment is a non-offshoring producer. Following this procedure, we obtain $\ln \hat{r}(1) = 10.640$ and can compute the mean of log revenues of non-

[20]Of course, the choice of moments is somewhat arbitrary. Therefore, we discuss in the appendix alternative moment conditions and the robustness of our results, when accounting for additional targets in the MM estimation.

offshoring establishments according to

$$\mu_2(\sigma) = 10.640 + (\sigma - 1)0.576. \tag{19}$$

The empirical counterpart to this moment, which we denote by m_2, can be directly observed in the dataset.

We use the two remaining moments to capture the overlap by targeting the share of offshoring establishments at two deciles of the revenue distribution. To link establishments with marginal costs c to their quantile position in the revenue (marginal cost) distribution, we have to compute the fraction of establishments with marginal costs lower than or equal to c: $G_c(c) = \int_0^c g_c(c)dc$. Accounting for Eq. (4), this allows us to compute the marginal cost corresponding to a certain quantile position in the revenue distribution according to $c_q = G_c^{-1}(q)$. With he marginal cost level at hand, the share of offshoring firms at quantile position q is then given by $\chi(c_q)$, according to Eq. (10). This gives the relevant moment from theory, $\mu_j(\nu_0, \nu_1, k, \kappa; q)$, with $j = 3, 4$ indicating the number of the moment and q referring to the quantile position of establishments with marginal cost c_q in the revenue distribution, which in the case of deciles can take a value between 1 and 9. In principle, we could use any pair of deciles to construct the two remaining moment conditions. Since we know from Figure 1 that the overlap reaches a maximum with an equal share of offshoring and non-offshoring establishments at a high decile position, we choose deciles 8 and 9 to target this peak.

When determining the observed counterparts for moments 3 and 4 – denoted by m_3 and m_4, respectively – we face the problem that each decile position covers only a small fraction of firms, which potentially makes the data moments vulnerable to outliers and at the same time challenges the idea that inverse variances provide suitable weights for these moment conditions. To avoid such problems, we construct the observed share of offshoring establishments at a decile position by averaging over an interval that includes the decile observation plus the ten nearest percentile neighbors, i.e. we associate the share of offshoring establishments at the eight decile with the average share over percentiles $\{75, ..., 80, ..., 85\}$ and proceed in the same way to construct the observed share of offshoring establishments at the ninth decile.[21]

[21] We have checked how changing the length of the percentile interval around the targeted decile affects our results and found our results to be robust in this respect. We decided for a relatively long interval, as this allows us to base the computation of the weights of these moment conditions on more observations.

3.3 Estimation results and model fit

Applying the MM estimator to our dataset gives the parameter values reported in the upper panel of Table 2. There, we see that ν_1 is larger than ν_0, indicating a rather strong dependency of the two technology parameters z and κ in our model. A closer look at the numerical solutions reveals that accounting for the dependencies of the two technology parameters is indeed crucial for our estimation, since there is not a single parameter combination with $\nu_1 = 0$ that fulfills the various constraints of our model. The value of σ is slightly lower than the one structurally estimated by Egger et al. (2013) when using firm-level information the IAB Establishment Panel. Our estimate of k is higher than the one reported by Egger et al. (2013), but is still in the range of shape parameters of the Pareto distribution from other studies (cf. Arkolakis, 2010; Arkolakis and Muendler, 2010).

Regarding the effective wage differential between the host and the source country of offshoring, κ, reliable estimates are not easy to find in the literature, mainly because reliable information on labor costs for a large sample of countries is not available. However, the US Bureau of Labor Statistics provides data on hourly labor compensation costs in manufacturing industries for a sample of 34 (mainly OECD member) countries over the period 1996-2012.[22] For some years, they provide information on labor compensation costs for two additional economies, namely India and China. Using information on bilateral trade from the OECD STAN Database, this allows us to construct a sample of 32 host countries of German offshoring, for which we have information on both the value of intermediate goods imports and the hourly labor compensation costs.[23] We use this information to construct an intermediate-goods-import-weighted measure of foreign labor compensation costs for the year of 2002, which amounts to 69.93 percent of the labor compensation costs reported for Germany in this year. Based on this value, we are confident, that our κ-estimate is of reasonable magnitude.

To draw statistical inference, we bootstrap the dataset 200 times and repeat for all these newly constructed datasets the simulation exercise outlined above. This establishes bootstrapped estimates for the parameters of interests, which are reported in Table 3, with standard errors in parentheses to the right of the parameter estimates and the 95 percent confidence

[22]Labor compensation costs cover all payments made directly to the worker, social insurance expenditures, and labor-related taxes (cf. http://www.bls.gov/fls/ichcc.pdf for further details).

[23]This country sample includes 14 of the 20 biggest suppliers of German intermediate goods imports and it covers, 84.13 of bilateral intermediate goods imports to Germany reported by the OECD STAN Database.

Table 2: Results for the baseline scenario

Parameter values					
	ν_0	ν_1	σ	k	κ
Estimates	0.12	0.88	5.80	7.23	0.75
Targeted moments					
	m_1	m_2	m_3	m_4	
Computed	0.01	13.41	0.50	0.62	
Observed	0.04	13.94	0.56	0.64	
Difference	-0.03	-0.53	-0.06	-0.02	

interval displayed below these estimates. We see that the results from the baseline scenario are close to and lie within the 95 percent confidence intervals of the parameter estimates from the bootstrapped data. In the case of ν_1, the confidence interval is fairly large, which may cast doubt on the preciseness of this estimate. Looking at the bootstrapped data in more detail, we find that only 21 out of the 200 ν_1-values are lower than 0.57, whereas 168 of the ν_1-levels lie within interval $[0.8, 0.9]$. In all cases, ν_1 takes a value larger than zero, which further substantiates that accounting for a dependency in the two technology parameters z and s is important for the parameter estimation.

Table 3: Estimation results from bootstrapping

Parameter values					
	ν_0	ν_1	σ	k	κ
Estimates	0.14 (0.00)	0.78 (0.02)	6.02 (0.03)	7.43 (0.05)	0.73 (0.00)
	[0.1, 0.22]	[0.05, 0.9]	[5.08, 6.45]	[5.93, 8.41]	[0.64, 0.77]
Targeted moments					
	m_1	m_2	m_3	m_4	
Computed	0.01 (0.00)	13.53 (0.01)	0.49 (0.00)	0.63 (0.00)	
Observed	0.03 (0.00)	13.94 (0.00)	0.56 (0.00)	0.64 (0.00)	
Difference	-0.02 (0.00)	-0.41 (0.01)	-0.07 (0.00)	-0.01 (0.00)	

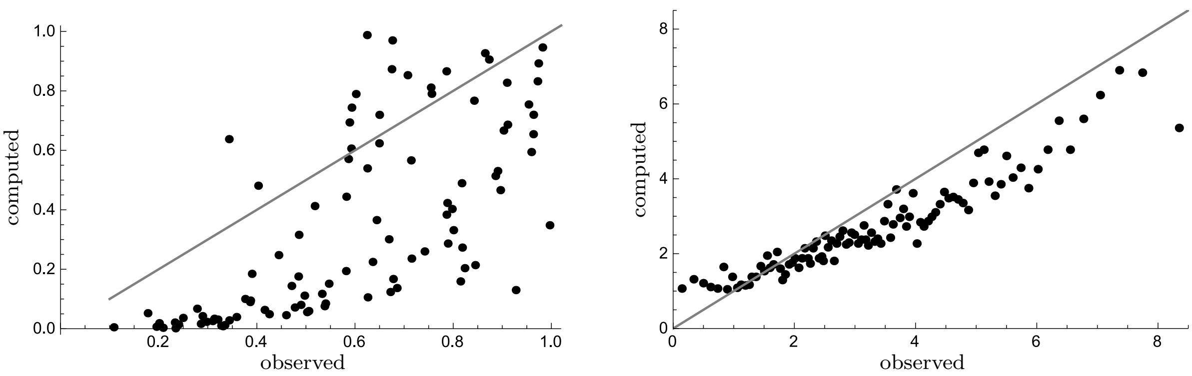
Note: Estimates refer to the mean value of 200 bootstraps. Standard errors in parentheses to the right of the estimates; 95 percent confidence interval reported below the parameter estimates.

Regarding the model fit, we can infer from the lower panel of Table 3 that the minimum distance estimator does a particularly good job in targeting moment conditions 2 and 4, as for these moments the computed values deviate by less than 10 percent from the observed ones. To shed further light on the aptitude of our model to capture important features of the data, we contrast computed and observed values of overlap, revenues, and marginal costs. Thereby, we rank firms according to their positions in the revenue distribution and rely on information at the percentile level. To obtain robust results, we construct the observed measures by averaging over an interval including the neighboring five millentiles above and below the respective percentile position.

Following this procedure, we can determine the share of offshoring establishments for each percentile in our dataset, $\widehat{\chi}_q$, and then construct the observed overlap for the various percentiles, according to $1 - |1 - 2\widehat{\chi}_q|$. This overlap measure is hump-shaped. It takes a value of 0 if either none or all establishments offshore and reaches a maximum value of 1 if the number of offshoring and non-offshoring firms is equal, i.e. if $\widehat{\chi}_q = 0.5$. The theory counterpart to this overlap measure can be constructed by computing the share of offshoring establishments for a certain quantile, $\chi(c_q)$, following the steps outlined in Section 3.2. The resulting value can then be used to compute a theory measure of overlap, according to $1 - |1 - 2\chi(c_q)|$. The fit between observed and computed overlap is reported in Panel A of Figure 4. There, we see that our model underestimates the overlap in the data. On average, the overlap in the model amounts to 0.35, which is about 58 percent of the observed overlap in the data. A closer look at the data reveals that our model underestimates the overlap in particular for low percentiles, whereas it captures the overlap nicely at higher percentiles of the revenue distribution. This is not surprising, as we target the share of offshoring firms at the eight and ninth decile in the MM estimation.

In Panel B of Figure 4, we contrast observed and computed values of log revenues of non-offshoring firms, which are normalized by subtracting the observed log revenue of the last productive establishment, $\ln r(1) = 10.64$.[24] The computed revenue values are constructed by means of Eq. (19), using the estimate of σ and the log of observed marginal costs (averaged over neighboring millentiles) as inputs. This exercise therefore provides insights regarding the explanatory power of the estimated σ, which according to the right Panel of Figure 4 is fairly good. However, the exercise does not shed light on the ability of our model to capture the distribution

[24] As outlined above, we cannot consider offshoring establishments here, as we lack detailed information on their marginal costs or the task content of their production process.

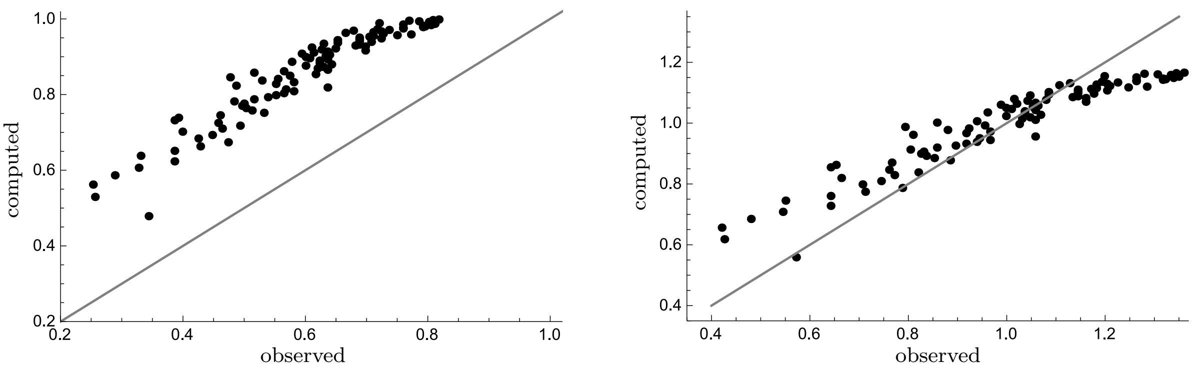


Panel A: Offshoring overlap

Panel B: Log normalized revenues

Figure 4: Model fit: Overlap and revenues

of marginal costs in the data, which has been used as input for computing the theory value of log-revenues. This is addressed by Figure 5, where we contrast observed and computed marginal costs of non-offshoring establishments at the percentile level. In Panel A, we see that our model systematically overestimates the marginal costs for all percentiles. In Panel B, we correct for the average upward bias by dividing marginal costs at each percentile by their mean value. This reveals that, although not overly successful in capturing the level of marginal costs, our model does a decent job in explaining the variance of marginal costs in the sample of non-offshoring establishments, which was one of the targeted moments in the MM estimation.



Panel A: Marginal costs

Panel B: Normalized marginal costs

Figure 5: Model fit: Marginal costs

3.4 Welfare effects of offshoring

We complete the discussion in this section, by employing the parameter estimates from Table 3 to quantify the welfare effects of offshoring in our model. For this purpose, we first note that welfare under autarky (superscript a) can be inferred from Eq. (16) by setting both $\hat{c}$ and κ equal to zero. The welfare effects of offshoring can then be computed according to $\Delta W = 100\,(W/W^a - 1)$:

$$\Delta W = 100 \left\{ \left(1 + \frac{\kappa L^*}{\tau L}\right)^{\frac{1}{\sigma-1}} \left[1 - \frac{\hat{c}^k}{\nu - \nu_1 \hat{c}} \left(\frac{\nu(\sigma-1)}{k-\sigma+1} - \frac{\nu_1 \hat{c}(\sigma-2)}{k-\sigma+2}\right) \frac{f}{f_e}\right]^{\frac{1}{1-\sigma}} - 1 \right\}. \qquad (20)$$

In a next step, we combine Eqs. (6), (15), and $\Gamma_2(\kappa, \hat{c}) = 0$ to solve for theory-consistent values of f, f_e, and $\tau L/L^*$ as functions of the five parameters ν_0, ν_1, σ, k, κ, and substitute the resulting expressions into Eq. (20). This gives the welfare effects of offshoring as function of the parameter estimates in Table 3. Following this approach, we find that according to our model offshoring has increased GDP per capita in Germany by 15.57 percent.

For an interpretation of this effect, the welfare estimate must be put in perspective to estimates reported by other studies. An interesting point of reference in this respect is the multi-country Ricardian trade model of Eaton and Kortum (2002), which has become an important workhorse of the quantitative trade literature in recent years. Eaton and Kortum compute the welfare effects of a country's movement from its observed trade openness to autarky and therefore consider a comparative static experiment of similar magnitude as ours. For Germany, they report a welfare loss of only 1.3 percent from moving towards autarky. Alvarez and Lucas (2007) use a Eaton and Kortum (2002)-type model in a calibration exercise for 60 economies. They provide a recipe on how to use their setting for computing an upper bound of the welfare gain associated with a movement from autarky to free trade. In the case of Germany the upper bound of welfare gain is 19.6 percent of GDP. However, Alvarez and Lucas point out that their upper bounds are unrealistically high numbers as they have been computed under the caveat of a frictionless open economy. In comparison with such findings, the welfare gains from offshoring may seem to be exaggerated by our model.

However, this would be a rash conclusion, because one has to keep in mind that the parsimonious structure of the quantitative models outlined above makes the welfare effects of trade (unrealistically) small. For instance, as pointed out by Caliendo and Parro (2015) welfare estimates become significantly larger in quantitative trade models when accounting for intermediate goods. Costinot and Rodriguez-Clare (2014) analyze the role of intermediates more systematically in their recent handbook article and show in an illustrative example that in the case of Germany, accounting for intermediates can lead to welfare estimates that are ten times higher than estimates from models that do not account for intermediates. We therefore think that the welfare estimates from our model are of reasonable magnitude.

4 The quantitative effect of ignoring the overlap in the data

Addressing the overlap of offshoring and non-offshoring establishments in the revenue distribution is the main purpose of this study. To shed light on how important it is to take this particular feature of the data into account, we consider a variant of our model without overlap. For this purpose, we set $\nu_0 = 1$ and $\nu_1 = 0$, implying that the probability of offshoring $\Pr_z(s > 0)$ equals one, when the fixed cost investment of f has been made. We can then estimate the two model parameters $\hat{c}$ and σ using the minimum distance estimator outlined in the previous section. The model constraints imposed by Eqs. (11) and (12) can be used to solve for theory-consistent values of k and κ, whereas conditions $k, \sigma > 1$ and $k > \sigma - 1$ confine the possible parameter space. As moment conditions for the minimization problem we consider the variance of the marginal costs of non-offshoring establishments as well as the link between revenues and marginal costs within these establishments established by Eq. (19). We choose this approach as our baseline specification to make the new estimates directly comparable to the parameter estimates in Table 3, where we have added the share of offshoring establishments at the eighth and ninth decile as two additional moment conditions to target the overlap. We estimate the model parameters using the same 200 bootstrapped datasets as in the previous section and report the results of this exercise in Table 4.

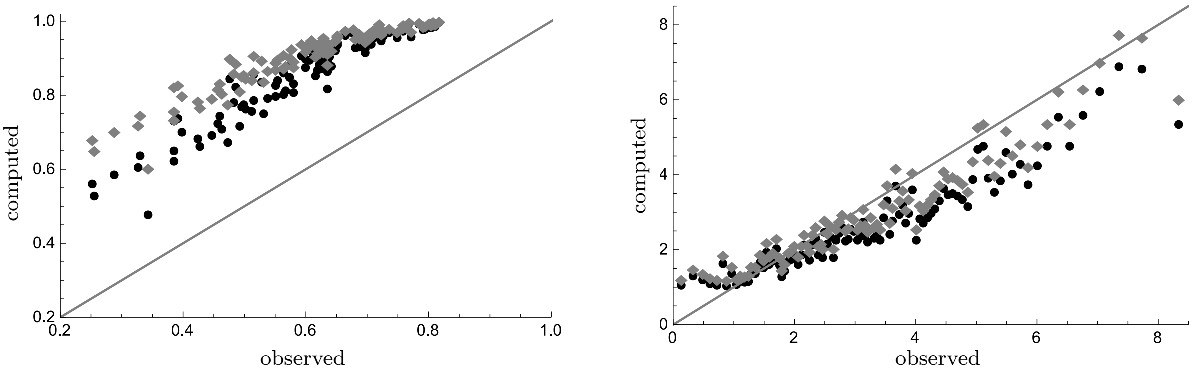
Compared to the parameter values in Table 3, we estimate higher values of σ, k, and a particularly (and perhaps unrealistically) high value of κ. The estimation results fulfill the moment conditions well, which is not surprising as the number of estimated parameters just equals the number of moment conditions. To assess the model fit, we display in Figure 6 computed against observed data points for marginal costs (Panel A) and revenues (Panel B) at the percentile level. We contrast the outcome for the model variant with overlap (black dots) with the outcome in the model variant without overlap (gray diamonds). From this Figure, we can conclude that ignoring the overlap somewhat decreases the fit of the model with the observed pattern of marginal costs and somewhat increases the fit of the model with the observed pattern of revenues. Overall, there seems to be not much difference between the two model variants regarding their fir with these to data moments. Of course, this conclusion is drawn under the

Table 4: Estimation results for the no-overlap case

Parameter values				
	$\hat{c}$	σ	k	κ
Estimates	0.85 (0.00)	6.63 (0.01)	9.10 (0.03)	0.997 (0.00)
	[0.84, 0.86]	[6.41, 6.82]	[8.49, 9.80]	[0.99, 1.00]
Targeted moments				
	m_1	m_2		
Computed	0.00 (0.00)	13.88 (0.00)		
Observed	0.03 (0.00)	13.94 (0.00)		
Difference	-0.03 (0.00)	-0.06 (0.00)		

Note: Estimates refer to the mean value of 200 bootstraps. Standard errors in parentheses to the right of the estimates; 95 percent confidence interval reported below the parameter estimates.

caveat that without overlap, we target less moment conditions and thus face fewer trade offs when estimating our parameters.



Panel A: Marginal costs

Panel B: Revenues

Figure 6: Model fit: observed vs. simulated data points

Using the parameter values in Table 4 we can conclude that according to the model without overlap offshoring has increased German GDP per capita by 5.12 percent. This is about 10.45 percentage points or 67.14 percent lower than in the model variant with overlap.[25] The reason

[25]Whereas the differences in welfare effects are sizable, we cannot be sure a priori that these differences are also significant from a statistical point of view. To shed light on the statistical significance, we can estimate the average welfare effects for the two model variants, using the bootstrapped datasets. This gives for the model variant with overlap an estimate of 15.38 with a 95 percent confidence interval of $[13.00, 19.41]$. For the model variant without overlap we estimate an average welfare effect of 5.13, with a 95 percent confidence interval of $[4.77, 5.47]$. The thus estimated welfare effects of offshoring in the model variants with and without overlap are

for this sizable gap lies in the difference of the κ-estimates in the two model variants, which reflects a fundamental bias from ignoring the overlap of offshoring and non-offshoring firms in quantitative offshoring models. Since the model without overlap associates offshoring with the most productive producers, it underestimates the (marginal) cost saving from offshoring. With the gains from offshoring being directly linked to its (marginal) cost saving effect, they are downward biased, when not accounting for the overlap in the data.

5 Offshoring at the turn of the millennium

With the parameter estimates from the previous sections at hand, we can use our model to dissect the German offshoring experience between 1990 and 2014. For this purpose, we first construct a comprehensive measure of offshoring that accounts for the import of both goods and service inputs, combining information from the OECD STAN database and the WITS database of the Worldbank. In the appendix, we provide details on how we construct our offshoring measure. Dividing the resulting measure by GDP gives the German openness to offshoring, e_{off}, which we target with our model in this section. The average openness observed in the data for the years 1998, 2000, and 2002, in which we know the offshoring status of German establishments, equals $\widehat{e}_{\text{off}} = 0.23$. A theory-consistent measure of offshoring openness can be computed according to $e_{\text{off}} = \kappa/(\tau L/L^* + \kappa)$, where $\tau L/L^*$ can be expressed as function of the parameter estimates in Table 3, according to Eqs. (6), (15), and $\Gamma_2(\kappa, \hat{c}) = 0$ (see above). For the model variants with and without overlap we compute an offshoring openness of 0.43 and 0.32, respectively. This implies that our model overestimates the average offshoring openness of Germany in the years 1998, 2000, and 2002, with the upward bias being even more pronounced for the model variant with overlap.

One may be tempted to conclude that the large discrepancy in the welfare effects of offshoring between the model variant with overlap and the model variant without overlap may at least partly be a result of the exaggeration of offshoring openness in the model. To see whether such a concern is justified, we can compute the welfare gain for a counterfactual situation, in which the offshoring openness in the model concurs with the data. We can do so, by adjusting the three parameters $\tau L/L^*$, κ, and $\hat{c}$ to align computed and observed values of e_{off} subject to the offshoring indifference and the labor market clearing conditions imposed by our model.

therefore statistically different at the 5 percent level.

The welfare gains of offshoring evaluated at the thus determined parameter values amount to 12.09 and 3.77 percent, respectively. Hence, our finding that the model variant without overlap significantly underestimates the welfare effects of offshoring is not the result of an exaggeration of offshoring openness by our model.

We now make use of our model to decompose the empirically observed evolution of German offshoring openness between 1990 and 2014 into changes at the intensive margin – capturing changes in the offshoring activity of incumbent offshoring firms – and the extensive margin – capturing changes in the mass of offshoring establishments. For this purpose, we compute for each year theory-consistent values of exogenous effective relative domestic labor supply $\tau L/L^*$ and the endogenous variables κ and $\hat{c}$ supporting the observed offshoring openness. We do these computations for both the model with overlap and the model without overlap and use the parameter estimates to determine theory-consistent values of offshoring openness for a counterfactual situation, in which the mass of offshoring establishments stays constant. We present a detailed analysis on how we compute these variables and an overview of our parameter estimates in the appendix and summarize the main insights from this decomposition exercise in the left Panel of Figure 7. The black line in this panel depicts the observed changes of German offshoring openness, whereas the solid and dashed gray lines capture the changes of offshoring that are attributed to the intensive margin by the model variants with and without overlap, respectively.

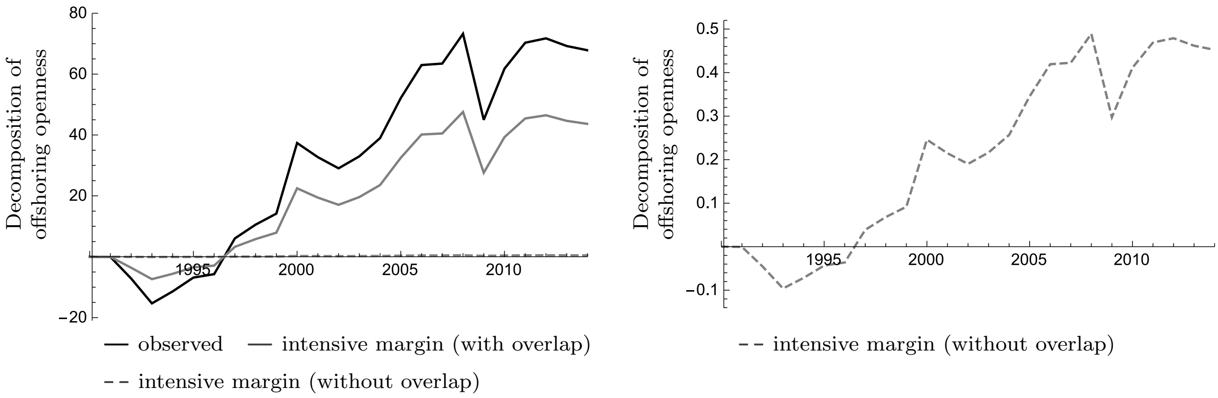


Figure 7: Changes in German offshoring openness between 1990 and 2014

The black line shows an overall increase in German offshoring openness since the early 1990s. However, this increase was not monotonic. There were ups and downs over this period, with three notable drops in the early 1990s, the early 2000s and, most strongly, in 2009. Despite a slight global decline in the trade to GDP ratio at the time, we think that the the first slump in offshoring openness reflects two particularities of the German reunification. Eastern German producers were less inclined to offshore, and western German producers gained access to cheap labor in the now larger domestic economy. The second slump picks up a general decline in the trade to GDP ratio in the aftermath of the dot-com crisis – maybe reinforced by a decline in the demand for cheap foreign labor after the drastic labor market reforms in Germany at the beginning of the new century. Finally, the drop of offshoring openness in 2009 captures the well documented decline in globalization during the financial crisis.

According to the model with overlap, both the extensive and intensive margin have plaid a prominent role in explaining the evolution of German offshoring openness, with the intensive margin contributing 64.34 percent to the overall increase in German offshoring exposure over the period 1990-2014. In contrast, the extensive margin is predominant in the model variant without overlap, explaining 99.34 percent of the observed increase in German offshoring exposure over the last 25 years. The left panel of Figure 7 may give the impression that offshoring did not change at all along the intensive margin. However, this is not true, as can be seen in the right panel of the figure, where we zoom in on the intensive margin of offshoring in the model variant without overlap. There, we see that the pattern of the intensive margin looks almost identical in the two model variants, but the range over which the intensive margin varies is significantly smaller in the model variant without overlap. This observation is well in line with the findings reported by Armenter and Koren (2015), who calibrate a quantitative trade model along the lines of Melitz (2003) with sorting of firms into export mode, using US data, and compare it with an otherwise identical trade model that allows for overlap of exporters and non-exporters. In a counterfactual exercise they show that lowering the iceberg trade cost parameter leads to substantial differences regarding the relative importance of the extensive and intensive margin for explaining the increase in exporting activity, with the extensive margin being more important in the model variant without overlap.

We complete the discussion in this section by simulating the gains from offshoring over the period 1990-2014. The results of this exercise are depicted by Figure 8. In line with our insights from Sections 3 and 4 the gains from offshoring are more pronounced when accounting for the observed overlap in the data. The welfare gain from the expansion of offshoring between 1990 and 2014 is 4.88 percent in the model variant with overlap and 1.80 percent in the model variant without overlap. To put the size of these effects into perspective, we can note that welfare in our model is a measure for per capita income, and hence we can contrast the offshoring gains with the overall increase in German GDP per capita between 1990 and 2014, which amounts to 38.55 percent. According to the model with overlap, the increased openness to offshoring has contributed 12.67 percent to the overall increase in German GDP per capita since 1990.

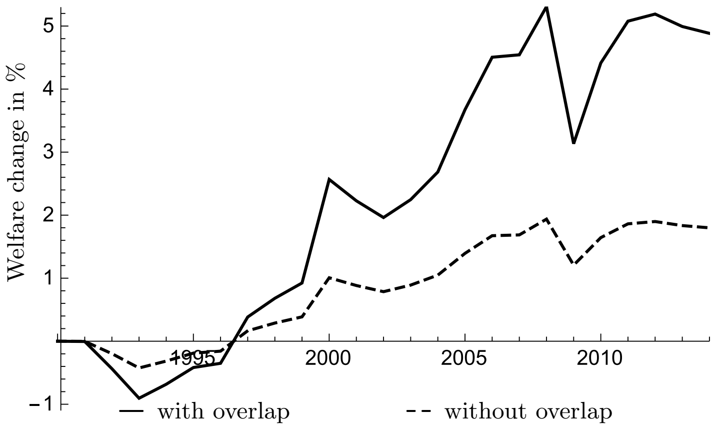


Figure 8: The welfare effects of Offshoring between 1990 and 2014

Taking stock, the analysis in this section confirms the previous result that ignoring the overlap leads to a severe downward bias in the welfare estimates of offshoring. Furthermore, the analysis shows that ignoring the overlap has the additional effect of exaggerating the contribution of the extensive margin to the observed increase of German offshoring openness between 1990 and 2014.

6 Robustness

The aim of this section is to study the robustness of our results, regarding the two main insights from our analysis: a downward bias in the welfare effects of offshoring and an exaggeration of the extensive margin of offshoring when ignoring the overlap in the data. For this purpose, we look at subsamples of the data – subsets of establishments and subsets of host countries – and and consider an alternative estimation strategy.[26]

[26]In the appendix, we present additional robustness checks, in which we consider additional moment conditions and account for an alternative classification of tasks in the MM estimation.

6.1 Subsets of establishments and host countries

In the first robustness check, we analyze whether pooling over industries is justified. To make sure that we are left with a sufficient amount of observations we do not use narrow definitions of industries, but instead distinguish between the two broad categories of manufacturing and services. We report the estimation results for the subsample of 7,581 establishment observations from manufacturing in Panels A (with overlap) and B (without overlap) of Table 5, whereas Panel C (with overlap) and D (without overlap) summarize the results for the subsample of 12,753 establishment observations from the service sector. Furthermore, since offshoring in our model is low-cost seeking, it is associated with production shifting from high-income to low-income countries. So far we have not distinguished host countries by their per-capita income levels, because the IAB Establishment Panel does not provide information on offshoring at the host country level. However, the dataset allows to distinguish offshoring to EMU members from offshoring to Non-EMU countries, so that we can to drop the former group from the sample of host countries. Whereas this does not exclude all high-income countries, it at least increases the relative frequency of low-income economies in the host country sample. We report the estimation results for the subsample of Non-EMU host countries in Panels E and F for the model variants with and without overlap, respectively.

Restricting the analysis to subsamples of establishments or host countries does affect our estimation results. For instance, we find that in 17 cases and thus 8.5 percent of the bootstrapped datasets for manufacturing $\hat{c}$ has a value of 1. In these cases, there is no selection into offshoring and all establishments make fixed cost investment f. Also, the estimated parameter values for the subsamples of establishments or host countries differ from those for the whole sample. Whereas the differences in the point estimates are in most cases not significant, they impact the quantitative effects of offshoring. The welfare gain from access to offshoring – i.e. from a movement from autarky to the observed extent of offshoring – in the model variant with overlap varies between 13.21 and 34.74 percent. Despite these quantitative differences, we find that in all three subsamples the downward bias in the welfare gain of offshoring when ignoring the overlap in the data is huge and varies between 68.14 percent (services) and 76.23 percent (manufacturing). Furthermore, according to the model with overlap the increase in German offshoring openness over the period 1990-2014 has been a strong welfare stimulus and associated with an increase in GDP per capita that ranges between 4.42 and 6.01 percent . The contribution

Table 5: Subsets of establishments and host countries

Panel A: Manufacturing with overlap

	ν_0	ν_1	σ	k	κ
Estimates	0.18 (0.00)	0.76 (0.01)	4.80 (0.05)	5.36 (0.09)	0.60 (0.00)
	ΔW		ΔW 1990-2014	Int. margin 1990-2014	
Effects	34.74%		6.01%		48.74%

Panel B: Manufacturing without overlap

	$\hat{c}$	σ	k	κ	
Estimates	0.83 (0.00)	5.66 (0.01)	6.22 (0.02)	0.998 (0.00)	
	ΔW		ΔW 1990-2014	Int. margin 1990-2014	
Effects	8.24%		2.15%		0.22%

Panel C: Services with overlap

	ν_0	ν_1	σ	k	κ
Estimates	0.21 (0.00)	0.18 (0.02)	6.72 (0.02)	6.89 (0.06)	0.74 (0.00)
	ΔW		ΔW 1990-2014	Int. margin 1990-2014	
Effects	13.21%		4.42%		74.37%

Panel D: Services without overlap

	$\hat{c}$	σ	k	κ	
Estimates	0.88 (0.00)	7.09 (0.01)	11.47 (0.06)	0.994 (0.00)	
	ΔW		ΔW 1990-2014	Int. margin 1990-2014	
Effects	4.21%		1.69%		1.25%

Panel E: Non-EMU with overlap

	ν_0	ν_1	σ	k	κ
Estimates	0.07 (0.00)	0.27 (0.02)	6.15 (0.03)	5.90 (0.05)	0.69 (0.01)
	ΔW		ΔW 1990-2014	Int. margin 1990-2014	
Effects	13.27%		5.55%		94.14%

Panel F: Non-EMU without overlap

	$\hat{c}$	σ	k	κ	
Estimates	0.79 (0.00)	6.58 (0.01)	9.35 (0.06)	0.999 (0.00)	
	ΔW		ΔW 1990-2014	Int. margin 1990-2014	
Effects	3.42%		1.80%		0.26%

Note: Estimates refer to the mean value of 200 bootstraps. Standard errors in parentheses to the right of the estimates.

of the intensive margin to the observed increase in offshoring exposure lies between 48.74 percent and 94.14, when accounting for the overlap. Both the welfare effects of offshoring as well as the contribution of the intensive margin to the overall increase in offshoring over the period 1990-2014 are considerably lower in the model variant without overlap.

Taking stock, the main result from the baseline model that ignoring the overlap leads to a severe downward bias in the welfare estimates of offshoring and to an exaggeration of the extensive margin of offshoring remains valid when distinguishing between manufacturing and services or dropping EMU members from the sample of host countries of offshoring.

6.2 An alternative estimation strategy

As outlined in Section 3.2, one can interpret the minimum-distance estimator in Eq. (17) as a GMM estimator with specific functional form of the moment conditions: $\Delta_t(\mathbf{o},\mathbf{x}) = m_t(\mathbf{o}) - \mu_t(\mathbf{x})$ for all $t = \{1, ..., n_m\}$. Using this interpretation, one may argue to use the inverse of the variance-covariance matrix of the moment conditions, $\mathbf{V}$, to construct weighting matrix $\mathbf{W}$: $\mathbf{W} = \mathbf{V}^{-1}$, as suggested by Hansen (1982) for GMM. Since $\mathbf{V}$ is not observable, we have to estimate it and do so by solving the minimum distance problem in Eq. (17) using the identity matrix for specifying a weighting matrix for a first-round estimation of parameters, which for the model variants with (without) overlap are summarized in Panel A (Panel B) of Table 6. We can use these parameter estimates to compute the four (two) theory moments $\mu_t(\cdot)$ for the model variant with (without) overlap and determine the difference between the thus computed moments and the data on the observational level – including only those observations that have been used for constructing data moment m_t. The difference between observations and computed moments gives observation-specific residuals, which we use to construct the variance-covariance matrix $\widehat{\mathbf{V}}$.

For instance, relying on the original dataset we obtain

$$\widehat{\mathbf{V}}_w = \begin{pmatrix} 0.04 & 0.81 & 0 & 0 \\ 0.81 & 17.06 & 0 & 0 \\ 0 & 0 & 0.22 & 0 \\ 0 & 0 & 0 & 0.25 \end{pmatrix}, \qquad \widehat{\mathbf{V}}_{w/o} = \begin{pmatrix} 0.06 & 0.90 \\ 0.90 & 17.07 \end{pmatrix} \tag{21}$$

for the model variants with and without overlap, respectively. The variance-covariance matrices are slightly different though structurally equivalent for the bootstrapped datasets. It is notable that the off-diagonal entries of $\widehat{\mathbf{V}}_w$ are zero in most cases. The reason is that the share of offshoring establishments at a decile position, which is constructed from observations in the neighborhood of the decile (see above), does not covary with other observations. On the one hand, the sets of establishment observations used for constructing the shares of offshoring firms at different deciles are disjoint. On the other hand, the residuals within the set of establishment observations used for constructing the share of offshoring firms at a certain decile are constant by construction if one looks only at non-offshoring firms. Hence, the respective residuals at the observational level do not covary with the residuals from the variance of marginal costs or the log of revenues.

We invert matrix $\widehat{\mathbf{V}}$ and check that the resulting weighting matrix $\mathbf{W}$ is positive semidefinite, and then use the thus determined weighting matrix in a second-round estimation of parameters. Panels C and D of Table 6 summarize the results of this two-step estimation approach for the model variants with and without overlap, respectively. There, we see that relying on the inverse of the variance-covariance matrix of moment conditions for specifying $\mathbf{W}$ increases the estimate of κ in the model variant with overlap. This lowers the welfare gain attributed to offshoring by this model variant by almost 50 percent. Furthermore, the welfare gain from the observed increase of German offshoring openness over the period 1990-2014 as well as the contribution of the intensive margin to this increase are also reduced drastically in the model variant with overlap. However, the main insights that ignoring the overlap leads to a downward bias in the welfare effects of offshoring and to an exaggeration of the extensive margin of offshoring remains valid when choosing an alternative weighting matrix.

7 Concluding remarks

This paper presents a model of heterogeneous firms, in which firms differ in the number of tasks they perform in the production process and the share of tasks they can offshore to a low-wage host country. Specific realizations of these two technology parameters are the outcome of a lottery and their distributions are interdependent. More specifically, we assume that firms, which perform more tasks, have a higher probability that at least some of their tasks are offshorable. Marginal production costs decline in the number of tasks performed and the share of tasks offshored. Offshoring is subject to fixed and variables costs, and not all firms find it attractive

Table 6: Alternative weighting of moment conditions

Panel A: $\mathbf{W}$ = identity matrix (with overlap)					
	ν_0	ν_1	σ	k	κ
Estimates	0.13 (0.00)	0.73 (0.01)	6.73 (0.01)	6.63 (0.02)	0.92 (0.00)
	ΔW		ΔW 1990-2014		Int. margin 1990-2014
Effects	8.03%		2.65%		18.44%

Panel B: $\mathbf{W}$ = identity matrix (without overlap)					
	$\hat{c}$	σ	k	κ	
Estimates	0.86 (0.00)	6.73 (0.01)	9.56 (0.03)	0.97 (0.00)	
	ΔW		ΔW 1990-2014		Int. margin 1990-2014
Effects	5.57%		1.98%		5.16%

Panel C: $\mathbf{W}$ = variance-covariance matrix of moment conditions (with overlap)					
	ν_0	ν_1	σ	k	κ
Estimates	0.14 (0.00)	0.61 (0.02)	6.06 (0.01)	5.68 (0.02)	0.90 (0.00)
	ΔW		ΔW 1990-2014		Int. margin 1990-2014
Effects	9.35%		3.15%		20.59%

Panel D: $\mathbf{W}$ = variance-covariance matrix of moment conditions (without overlap)					
	$\hat{c}$	σ	k	κ	
Estimates	0.83 (0.00)	5.80 (0.01)	7.92 (0.02)	0.98 (0.00)	
	ΔW		ΔW 1990-2014		Int. margin 1990-2014
Effects	6.27%		2.30%		3.86%

Note: Estimates refer to the mean value of 200 bootstraps. Standard errors in parentheses to the right of the estimates.

to make the investment into offshoring. This gives a model of heterogeneoues firms, in which some but not all firms of a certain (cost or) revenue category conduct offshoring, with the share of offshoring firms increasing in revenues.

In an empirical exercise, we estimate key parameters of the model, using firm-level data from Germany. Thereby, we rely on a method of moments approach, in which we estimate the parameters of our model by minimizing the distance between moments observed in the data and their theoretical counterparts computed from the model. Based on the parameter estimates, we show that access to offshoring has increased welfare in Germany by 12.67 percent. This welfare estimate is more than 67 percent higher than in an otherwise identical model without overlap, which is intuitive, because in a model without overlap offshoring is associated with high-productivity firms and thus with firms, which by assumption require just a small marginal cost saving for finding production shifting to a low-wage country attractive. Furthermore, in a decomposition analysis we show that the increase in German offshoring over the period 1990-2014 was to a large extent driven by an increase along the intensive margin, i.e. by an expansion in offshoring of incumbent offshoring firms. This differs from the decomposition obtained in a model that ignores the overlap of offshoring and non-offshoring firms in the data. Such a model would attribute the increase in German offshoring openness over the respective time period almost entirely to the extensive margin, i.e. to the increase in the number of offshoring firms. We show that the two main insights of a downward bias in the welfare effects of offshoring and the exaggeration of the extensive margin of offshoring when ignoring the overlap in the data are robust to changes in the composition of firms or host countries and are also robust to changes in the estimation strategy.

Whereas elaborating on two important biases that materialize when ignoring the overlap of offshoring and non-offshoring firms in the data can provide an important stimulus for future research on the quantitative effects of offshoring, the parsimonious structure – helpful for analytical tractability and for illustrative purposes – limits the aptitude of our model to capture important features of offshoring in the real world. For instance, by associating the host country with a labor reservoir whose population lacks entrepreneurial abilities, our model cannot address the crowding out of local production by foreign labor demand of offshoring firms in the host country. Accordingly, our model is not well suited for studying the welfare effects in the host country of offshoring and for analyzing to what extent these welfare estimates are biased when ignoring the overlap in the data. Extending the model to one that allows for firms also in the host country of offshoring therefore seems a promising avenue for future research.

References

Acemoglu, D. and D. Autor (2011): "Skills, Tasks and Technologies: Implications for Employment and Earnings," in *Handbook of Labor Economics*, ed. by O. Ashenfelter and D. Card, North Holland/NL: Elsevier, vol. 4, 1043–1171.

Acemoglu, D., G. Gancia, and F. Zilibotti (2015): "Offshoring and Directed Technical Change," *American Economic Journal: Macroeconomics*, 7, 84–112.

Alvarez, F. and J. Lucas, Robert E. (2007): "General Equilibrium Analysis of the Eaton-Kortum Model of International Trade." *Journal of Monetary Economics*, 54, 1726 – 1768.

Antràs, P., L. Garicano, and E. Rossi-Hansberg (2006): "Offshoring in a Knowledge Economy," *Quarterly Journal of Economics*, 121, 31–77.

Antràs, P. and E. Helpman (2004): "Global Sourcing," *Journal of Political Economy*, 112, 552–580.

Arkolakis, C. (2010): "Market Penetration Costs and the New Consumers Margin in International Trade," *Journal of Political Economy*, 118, 1151–1199.

Arkolakis, C. and M.-A. Muendler (2010): "The Extensive Margin of Exporting Products: A Firm-level Analysis," NBER Working Papers 16641.

Armenter, R. and M. Koren (2015): "Economies of Scale and the Size of Exporters." *Journal of the European Economic Association*, 13, 482 – 511.

Becker, S. O., K. Ekholm, and M.-A. Muendler (2013): "Offshoring and the Onshore Composition of Tasks and Skills," *Journal of International Economics*, 90, 91–106.

Becker, S. O. and M.-A. Muendler (2015): "Trade and Tasks: An Exploration over Three Decades in Germany," *Economic Policy*, 30, 589–641.

Bergin, P. R., R. C. Feenstra, and G. H. Hanson (2011): "Volatility Due to Offshoring: Theory and Evidence," *Journal of International Economics*, 85, 163–173.

Bernard, A. B., J. B. Jensen, S. J. Redding, and P. K. Schott (2012): "The Empirics of Firm Heterogeneity and International Trade," *Annual Review of Economics*, 4, 283–313.

Blinder, A. S. (2006): "Offshoring: The Next Industrial Revolution?" *Foreign Affairs*, 85, 113–128.

Blinder, A. S. and A. B. Krueger (2013): "Alternative Measures of Offshorability: A Survey Approach," *Journal of Labor Economics*, 31, S97–S128.

Caliendo, L. and F. Parro (2015): "Estimates of the Trade and Welfare Effects of NAFTA." *Review of Economic Studies*, 82, 1 – 44.

Cameron, A. C. and P. K. Trivedi (2005): *Microeconometrics – Methods and Applications*, Cambridge University Press.

Costinot, A. and A. Rodriguez-Clare (2014): "Trade Theory with Numbers: Quantifying the Consequences of Globalization," in *Handbook of International Economics*, ed. by G. Gopinath, E. Helpman, and K. Rogoff, North Holland, vol. 4, chap. 4, 197–261.

Das, S., M. J. Roberts, and J. R. Tybout (2007): "Market Entry Costs, Producer Heterogeneity, and Export Dynamics." *Econometrica*, 75, 837 – 873.

Davis, D. R. and J. Harrigan (2011): "Good Jobs, Bad Jobs, and Trade Liberalization," *Journal of International Economics*, 84, 26–36.

Eaton, J. and S. Kortum (2002): "Technology, Geography, and Trade," *Econometrica*, 70, 1741–1779.

Eaton, J., S. Kortum, and F. Kramarz (2011): "An Anatomy of International Trade: Evidence from French Firms." *Econometrica*, 79, 1453 – 1498.

Egger, H., P. Egger, and U. Kreickemeier (2013): "Trade, Wages, and Profits," *European Economic Review*, 64, 332 – 350.

Egger, H., U. Kreickemeier, and J. Wrona (2015): "Offshoring Domestic Jobs," *Journal og International Economics*, 97, 112–125.

——— (2016): "Offshoring, Firm Selection, and Job Polarisation in General Equilibrium," Tech. rep., Unpublished Manuscript, University of Bayreuth.

Ferguson, T. S. (1958): "A Method of Generating Best Asymptotically Normal Estimates with Application to the Estimation of Bacterial Densities," *Annals of Mathematical Statistics*, 29, 1046–1062.

Fischer, G., F. Janik, D. Müller, and A. Schmucker (2008): "The IAB Establishment Panel – from Sample to Survey to Projection," *FDZ Methodenreport – Methodological aspects of labour market dataï¿½*, 01/2008.

Ghironi, F. and M. J. Melitz (2005): "International Trade and Macroeconomic Dynamics with Heterogeneous Firms." *Quarterly Journal of Economics*, 120, 865 – 915.

Grossman, G. M. and E. Rossi-Hansberg (2008): "Trading Tasks: A Simple Theory of Offshoring," *American Economic Review*, 98, 1978–1997.

Hall, A. R. (2005): *Generalized Methods of Moments*, Oxford University Press.

Hallak, J. C. and J. Sivadasan (2013): "Product and Process Productivity: Implications for Quality Choice and Conditional Exporter Premia." *Journal of International Economics*, 91, 53 – 67.

Hansen, L. P. (1982): "Large Sample Properties of Generalized Method of Moments Estimators," *Econometrica*, 50, 1029–1054.

Harrigan, J. and A. Reshef (2015): "Skill Biased Heterogeneous Firms, Trade Liberalization, and the Skill Premium," *Canadian Journal of Economics*, 48, 1024–1066.

Helpman, E., O. Itskhoki, M.-A. Muendler, and S. J. Redding (2016): "Trade and Inequality: From Theory to Estimation," *Revidew of Economic Studies*, forthcoming.

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

Levy, F. and R. J. Murnane (2004): *The New Division of Labor*, Princeton/NJ.: Princeton University Press.

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

Moser, C., D. M. Urban, and B. Weder di Mauro (2015): "On the Heterogeneous Employment Effects of Offshoring – Identifying Productivity and Downsizing Effects," *Economic Inquiry*, 53, 220–239.

Rodriguez-Clare, A. (2010): "Offshoring in a Ricardian World," *American Economic Journal: Macroeconomics*, 2, 227–258.

Rodriguez-Lopez, A. (2014): "What Drives Offshoring Decisions? Selection and Escape-Competition Mechanisms," Tech. rep., Unpublished Manuscript, University of California, Irvine.

Rohrbach-Schmidt, D. (2009): "The BIBB/IAB- and BIBB/BAuA Surveys of the Working Population on Qualification and Working Conditions in Germany," BIBB-FDZ Daten- und Methodenberichte 1/2009, Bundesinstitut für Berufsbildung Forschungsdatenzentrum.

Spitz-Oener, A. (2006): "Technical Change, Job Tasks and Rising Educational Demands: Looking Outside the Wage Structure," *Journal of Labor Economics*, 24, 235–270.

Windmeijer, F. (2004): "A Finite Sample Correction For the Variance of Linear Efficient Two-Step GMM Estimators," *Journal of Econometrics*, 126, 25–51.

A Theoretical appendix

A.1 Derivation of Eq. (4)

Let us define $b = 1 - z$. Then, the cumulative distribution of b is given by $G_b(b)$, with $g_b(b) = kb^{k-1}$. Since $c = b$ if $z \leq \hat{z}$ and thus $c \geq \hat{c}$, the third segment of the pdf of c is given by $g_c(c) = kc^{k-1}$. To determine the pdf of c for interval $c \leq \hat{c}$, we can note that only a share $\nu_0 + \nu_1 z$ of firms that make the fixed cost investment will end up to offshore. Substituting z by $1 - b$, the pdf of offshoring and non-offshoring firms is therefore given by $g_b^o(b) \equiv kb^{k-1}(\nu - \nu_1 b)$ and $g_b^d(b) \equiv kb^{k-1}(1-\nu+\nu_1 b)$, with $g_b(b) = g_b^d(b)+g_b^o(b) = kb^{k-1}$. For purely domestic producers, we have $c = b$, and can thus write $g_c^d(c) = kc^{k-1}(1 - \nu + \nu_1 c)$. For offshoring firms, things are different, because $\kappa < 1$ establishes $c = b\kappa^s < b$. Accounting for $a \equiv \kappa^s$, we can compute $s = \ln a / \ln \kappa$, and hence can write $\Pr(a \leq \bar{a}) = 1 - \Pr\big(s \leq s(\bar{a})\big) = 1 - \int_0^{s(\bar{a})} 1 ds = 1 - \ln \bar{a} / \ln \kappa$. The pdf of a can therefore be expressed as $g_a(a) = -1/(a \ln \kappa)$.

We can now compute the pdf of c for those firms that actually offshore, according to

$$g_c^o(c) = \int_{b \in B} g_b^2(b) g_a\left(\frac{c}{b}\right)\left|\frac{1}{b}\right| db = -\frac{1}{c \ln \kappa}\int_{b \in B} kb^{k-1}(\nu - \nu_1 b) db, \qquad \text{(A.1)}$$

where B is the set of feasible b's. To determine the bounds of the integral, we can note that b varies over the interval $[c, c/\kappa)$ if $c < \kappa\hat{b}$, whereas b varies over the interval $[c, \hat{b}]$ if $c \geq \kappa\hat{b}$. Let us first consider parameter domain $c < \kappa\hat{b}$. In this case, we have

$$
\begin{aligned}
g_c^{o1}(c) &= -\frac{k}{c\ln\kappa}\int_c^{c/\kappa}\left[\nu b^{k-1} - \nu_1 b^k\right] db \\
&= -\frac{1}{\ln\kappa}\left\{\nu c^{k-1}\left[\left(\frac{1}{\kappa}\right)^k - 1\right] - \nu_1\frac{kc^k}{k+1}\left[\left(\frac{1}{\kappa}\right)^{k+1} - 1\right]\right\}. \qquad \text{(A.2)}
\end{aligned}
$$

In contrast, if $c \geq \kappa\hat{b}$, we obtain

$$
\begin{aligned}
g_c^{o2}(c) &= -\frac{k}{c\ln\kappa}\int_c^{\hat{b}}\left[\nu b^{k-1} - \nu_1 b^k\right] db \\
&= -\frac{1}{\ln\kappa}\left\{\nu c^{k-1}\left[\left(\frac{\hat{b}}{c}\right)^k - 1\right] - \nu_1\frac{kc^k}{k+1}\left[\left(\frac{\hat{b}}{c}\right)^{k+1} - 1\right]\right\}. \qquad \text{(A.3)}
\end{aligned}
$$

Replacing $\hat{b}$ by $\hat{c}$ and adding up $g_c^d(c)$ and $g_c^o(c)$ for the two parameter domains gives the first and the second segment of the probability density function in Eq. (4). This completes the proof. *QED*

A.2 Derivation of Eq. (7)

Accounting for $r(c)/r(1) = c^{1-\sigma}$, aggregate revenues can be written as $R = M\int_0^1 r(c)g_c(c)dc = Mr(1)\int_0^1 c^{1-\sigma}g_c(c)dc$. We have to compute the integrals separately for the three segments of $g_c(c)$. For the first segment, we obtain

$$
\begin{aligned}
R_1 &= Mr(1)\int_0^{\kappa\hat{c}} c^{1-\sigma}g_c(c)dc \\
&= Mr(1)\int_0^{\kappa\hat{c}} c^{1-\sigma}\Bigg\{(1-\nu+\nu_1 c)kc^{k-1} \\
&\quad - \frac{1}{\ln\kappa}\left\{\nu c^{k-1}\left[\left(\frac{1}{\kappa}\right)^k - 1\right] - \nu_1\frac{kc^k}{k+1}\left[\left(\frac{1}{\kappa}\right)^{k+1} - 1\right]\right\}\Bigg\} dc. \qquad \text{(A.4)}
\end{aligned}
$$

Solving for the integral, gives

$$
\begin{aligned}
R_1 = Mr(1)\Bigg\{&(1-\nu)\frac{k}{k-\sigma+1}(\kappa\hat{c})^{k-\sigma+1} + \nu_1\frac{k}{k-\sigma+2}(\kappa\hat{c})^{k-\sigma+2} \\
&-\nu\frac{\kappa^{-k}-1}{\ln[\kappa]}\frac{1}{k-\sigma+1}(\kappa\hat{c})^{k-\sigma+1} + \nu_1\frac{\kappa^{-(k+1)}-1}{\ln[\kappa]}\frac{k}{k+1}\frac{1}{k-\sigma+2}(\kappa\hat{c})^{k-\sigma+2}\Bigg\}. \qquad \text{(A.5)}
\end{aligned}
$$

Thereby, $k > \sigma - 1$ has been assumed to obtain a finite value of R_1. For the second segment, we can write

$$
\begin{aligned}
R_2 &= Mr(1) \int_{\kappa\hat{c}}^{\hat{c}} c^{1-\sigma} g_c(c) dc \\
&= Mr(1) \int_{\kappa\hat{c}}^{\hat{c}} c^{1-\sigma} \Bigg\{ (1 - \nu + \nu_1 c) k c^{k-1} \\
&\quad - \frac{1}{\ln \kappa} \left\{ \nu c^{k-1} \left[\left(\frac{\hat{c}}{c} \right)^k - 1 \right] - \nu_1 \frac{k c^k}{k+1} \left[\left(\frac{\hat{c}}{c} \right)^{k+1} - 1 \right] \right\} \Bigg\} dc. \qquad \text{(A.6)}
\end{aligned}
$$

Solving for the integral establishes

$$
\begin{aligned}
R_2 = Mr(1) \Bigg\{ & (1-\nu) \frac{k}{k-\sigma+1} \left(1 - \kappa^{k-\sigma+1}\right) \hat{c}^{k-\sigma+1} + \nu_1 \left(1 - \kappa^{k-\sigma+2}\right) \frac{k \hat{c}^{k-\sigma+2}}{k-\sigma+2} \\
& + \nu \frac{1-\kappa^{1-\sigma}}{\ln[\kappa]} \frac{\hat{c}^{k-\sigma+1}}{\sigma-1} + \nu \frac{1-\kappa^{k-\sigma+1}}{\ln[\kappa]} \frac{\hat{c}^{k-\sigma+1}}{k-\sigma+1} \\
& - \nu_1 \frac{1-\kappa^{1-\sigma}}{\ln[\kappa]} \frac{k \hat{c}^{k-\sigma+2}}{(\sigma-1)(k+1)} - \nu_1 \frac{1-\kappa^{k-\sigma+2}}{\ln[\kappa]} \frac{k \hat{c}^{k-\sigma+2}}{(k+1)(k-\sigma+2)} \Bigg\}. \qquad \text{(A.7)}
\end{aligned}
$$

Finally, for the third segment, we obtain

$$
\begin{aligned}
R_3 &= Mr(1) \int_{\hat{c}}^{1} c^{1-\sigma} g_c(c) dc = Mr(1) \int_{\hat{c}}^{1} c^{1-\sigma} k c^{k-1} dc \\
&= Mr(1) \left[\frac{k}{k-\sigma+1} - \frac{k}{k-\sigma+1} \hat{c}^{k-\sigma+1} \right]. \qquad \text{(A.8)}
\end{aligned}
$$

Total revenues in Eq. (7) can then be computed by adding up R_1, R_2 and R_3. This completes the proof. *QED*

A.3 Properties of the offshoring indifference condition

Let us define

$$
\alpha(\kappa) = \frac{\kappa^{1-\sigma} - 1}{(1-\sigma)\ln\kappa} - 1, \qquad \text{(A.9)}
$$

with $\alpha(0) = \lim_{\kappa \to 0} \kappa^{1-\sigma} - 1 = \infty$, $\alpha(1) = \lim_{\kappa \to 1} \kappa^{1-\sigma} - 1 = 0$, and

$$
\alpha'(\kappa) = \frac{\hat{\alpha}(\kappa)}{(1-\sigma)[\ln\kappa]^2 \kappa^{\sigma}}, \qquad \hat{\alpha}(\kappa) \equiv (1-\sigma)\ln\kappa + \kappa^{\sigma-1} - 1. \qquad \text{(A.10)}
$$

Accounting for $\lim_{\kappa \to 0} \hat{\alpha}(\kappa) = \infty$, $\hat{\alpha}(1) = 0$, and $\hat{\alpha}'(\kappa) = [(\sigma-1)/\kappa](\kappa^{\sigma-1} - 1) < 0$, it follows that $\hat{\alpha}(\kappa) > 0$ holds for all possible $\kappa < 1$. Considering $\sigma > 1$, we get $\alpha'(\kappa) < 0$. This allows us to compute

$$\frac{\partial \Gamma_1(\cdot)}{\partial \kappa} = \left\{ \frac{\hat{c}^k}{\nu - \nu_1 \hat{c}} \left[\nu \frac{\sigma - 1}{k - \sigma + 1} - \nu_1 \hat{c} \frac{\sigma - 2}{k - \sigma + 2} \right] - \frac{f_e}{f} \right\} \alpha'(\kappa) \tag{A.11}$$

and, since the bracket expression must be negative if $\Gamma(\cdot) = 0$, we can safely conclude that $\partial \Gamma_1(\cdot)/\partial \kappa > 0$.

Differentiation $\Gamma_1(\cdot)$ with respect to $\hat{c}$ yields

$$\frac{\partial \Gamma_1(\cdot)}{\partial \hat{c}} = \frac{\hat{c}^{\sigma-1}}{\nu - \nu_1 \hat{c}} \frac{k}{k - \sigma + 1} \left(\frac{\sigma - 1}{\hat{c}} + \frac{\nu_1}{\nu - \nu_1 \hat{c}} \right) + \frac{k \hat{c}^{k-1}}{\nu - \nu_1 \hat{c}} \left[\frac{(\sigma - 1)\nu}{k - \sigma + 1} - \frac{(\sigma - 2)\nu_1 \hat{c}}{k - \sigma + 2} \right] \alpha(\kappa)$$
$$+ \frac{\nu_1 \hat{c}^k}{(\nu - \nu_1 \hat{c})^2} \left[\frac{(\sigma - 1)\nu}{k - \sigma + 1} - \frac{(\sigma - 2)\nu}{k - \sigma + 2} \right] \alpha(\kappa).$$

In view of $\hat{c} \leq 1$, the first two expressions on the right-hand side of this derivative must be positive. Furthermore, it follows from $(\sigma - 1)(k - \sigma + 2) > (\sigma - 2)(k - \sigma + 1)$ that the third term is positive as well. This implies $\partial \Gamma_1(\cdot)/\partial \hat{c} > 0$. Putting together, we can conclude that if $\Gamma_1(\cdot) = 0$ has an interior solution in $\hat{c}$ and κ, the relationship between $\hat{c}$ and κ established by $\Gamma_1(\cdot) = 0$, can be determined according to

$$\left. \frac{d\hat{c}}{d\kappa} \right|_{\Gamma_1(\cdot)=0} = - \frac{\partial \Gamma_1/\partial \kappa}{\partial \Gamma_1/\partial \hat{c}} \tag{A.12}$$

and is negative.

To see for which domains of $\hat{c}$ and κ $\Gamma_1(\hat{c}, \kappa) = 0$ is feasible, we can first note that if κ reaches a maximum level of one, $\alpha(\kappa)$ is zero, and hence $\hat{c}$ must fall to zero in order to restore $\Gamma_1(\hat{c}, \kappa) = 0$. In contrast, two cases have to be distinguished regarding the minimum level of κ and the maximum level of $\hat{c}$. If $\Gamma_1(\hat{c}, 0) = 0$ has a solution in $\hat{c}$ on the unit interval, the minimum possible value of κ is equal to zero and the maximum possible value of $\hat{c}$ is smaller than (or equal to) one. If however $\Gamma_1(\hat{c}, 0) = 0$ has no solution with $\hat{c} \leq 1$, the maximum level of $\hat{c}$ is equal to one, whereas the minimum level of κ is larger than zero and implicitly determined by $\Gamma_1(1, \kappa) = 0$. A sufficiently high level of f ensures that the minimum possible κ is zero and that the maximum possible $\hat{c}$ is smaller than one. In this case, $\hat{c}$ is well defined as decreasing function of κ over the whole unit interval. This completes the formal discussion on the properties of OC. *QED*

A.4 Derivation of Eq. (8)

The foreign wage bill is given by

$$w^* L^* = M \frac{\sigma - 1}{\sigma} \int_0^{\hat{c}} \left[\int_0^1 s \hat{r}(b, s) ds \right] g_b^o(b) db, \tag{A.13}$$

with $\hat{r}(b,s) \equiv r(b\kappa^s) = r(c)$. Using the insight that $r(c)/r(1) = c^{1-\sigma}$, we thus obtain

$$w^* L^* = Mr(1)\frac{\sigma-1}{\sigma}\mathbb{E}\left[s\kappa^{s(1-\sigma)}\right]\int_0^{\hat{c}} b^{1-\sigma} g_b^o(b)db. \tag{A.14}$$

Substituting $g_b^o(b) = kb^{k-1}[\nu - \nu_1 b]$, we get

$$\int_0^{\hat{c}} b^{1-\sigma} g_b^o(b)db = \frac{k}{k-\sigma+1}\hat{c}^{k-\sigma+1}\left(\nu - \nu_1 \hat{c}\frac{k-\sigma+1}{k-\sigma+2}\right). \tag{A.15}$$

Furthermore, we can compute

$$\mathbb{E}\left[s\kappa^{s(1-\sigma)}\right] = \int_0^1 s\kappa^{s(1-\sigma)}ds = \frac{1+\kappa^{1-\sigma}[(1-\sigma)\ln\kappa - 1]}{[(1-\sigma)\ln\kappa]^2}. \tag{A.16}$$

Putting together, we thus obtain

$$w^* L^* = Mr(1)\frac{\sigma-1}{\sigma}\frac{k}{k-\sigma+1}\hat{c}^{k-\sigma+1}\left(\nu - \nu_1 \hat{c}\frac{k-\sigma+1}{k-\sigma+2}\right)\frac{1+\kappa^{1-\sigma}[(1-\sigma)\ln\kappa - 1]}{[(1-\sigma)\ln\kappa]^2}, \tag{A.17}$$

which, in view of Eq. (7), establishes Eq. (8). *QED*

A.5 Properties of the labor market constraint

Let us define $a \equiv \kappa^{1-\sigma} > 1$ and $b \equiv \left[\hat{c}^{k-\sigma+1}\left(\nu - \nu_1 \hat{c}\frac{k-\sigma+1}{k-\sigma+2}\right)\right]^{-1} > 1$. This allows us to rewrite Γ_2 as follows: $\Gamma_2 = \kappa\{[\sigma/(\sigma-1)]\beta(a,b) - 1\} - \tau L/L^*$, with

$$\beta(a,b) \equiv \frac{(b-1)[\ln a]^2 + (a-1)\ln a}{1 + a(\ln a - 1)}. \tag{A.18}$$

Differentiating $\beta(a,b)$ with respect to a, we find that $\beta_a'(a,b) >,=,< 0$ is equivalent to $\hat{\beta}(a,b)\ln a > ,=,< 0$, with

$$\hat{\beta}(a,b) \equiv 2(b-1)\ln a - 2(b-1)\frac{a-1}{a} - (b-1)[\ln a]^2 + \ln a - \frac{(a-1)^2}{a\ln a} \tag{A.19}$$

Partially differentiating $\hat{\beta}(a,b)$ with respect to a, establishes

$$\hat{\beta}_a'(a,b) = 2\frac{b-1}{a^2}(a - 1 - a\ln a) - \frac{(a^2-1)\ln a - (a-1)^2 - a[\ln a]^2}{[a\ln a]^2}. \tag{A.20}$$

The first term on the right-hand side is unambiguously negative for all $a > 1$. To determine the sign of the second term, we can differentiate $B(a) \equiv (a^2-1)\ln a - (a-1)^2 - a[\ln a]^2$, which gives $B'(a) = \hat{B}(a)/a$, with $\hat{B}(a) \equiv 2a(a-1)\ln a - (a-1)^2 - a[\ln a]^2$. Noting that $\hat{B}(1) = B(1) = 0$ and that $\hat{B}'(a) = \ln a[4(a-1) - \ln a] > 0$ holds for all $a > 1$, it follows that $B(a)$ must be positive, and we can thus safely conclude that $\hat{\beta}_a'(a,b) < 0$. In view of $\hat{\beta}(1,b) = 0$, this establishes

$\beta_a'(a,b) < 0$, which, in view of $da/d\kappa < 0$, is sufficient for $d\Gamma/d\kappa|_{\Gamma_2=0} > 0$. Furthermore, it is immediate that $\beta_b'(a,b) > 0$. Accounting for

$$\frac{db}{d\hat{c}} = -\frac{(k-\sigma+1)\hat{c}^{k-\sigma}(\nu-\nu_1\hat{c})}{\left[\hat{c}^{k-\sigma+1}\left(\nu-\nu_1\hat{c}^{\frac{k-\sigma+1}{k-\sigma+2}}\right)\right]^2} < 0, \tag{A.21}$$

we thus obtain $d\Gamma/d\hat{c}|_{\Gamma_2=0} < 0$. Putting together, the implicit function theorem establishes $d\kappa/d\hat{c}|_{\Gamma_2=0} > 0$, provided that $\Gamma_2(\kappa;\hat{c}) = 0$ has a solution in the $(\kappa,\hat{c})$-space.

To see for which domains of κ and $\hat{c}$ $\Gamma_2(\kappa,\hat{c}) = 0$ is feasible, we can first note that if $\hat{c}$ falls to a minimum level of 0, κ must go to zero as well in order to restore $\Gamma_2(\kappa,\hat{c}) = 0$. In contrast, two cases must be distinguished regarding the maximum levels of $\hat{c}$ and κ. If for $\hat{c} = 1$ $\Gamma_2(\kappa,1) = 0$ has a solution in κ on the unit interval, the maximum possible value of κ is smaller than (or equal to) one, whereas the maximum possible value of $\hat{c}$ is one. If however $\Gamma_2(\kappa,1) = 0$ has no solution for $\kappa \leq 1$, then the maximum possible value of κ is equal to one, whereas the maximum possible value of $\hat{c}$ is smaller than one and implicitly determined by $\Gamma_2(1,\hat{c}) = 0$. A sufficiently high level of τ ensures that the first case is realized, implying that κ is well defined as increasing function of $\hat{c}$ over the whole unit interval This completes the formal discussion on the properties of LC. *QED*

A.6 The impact of changes in τ and f on W

Totally differentiating Eq. (16) with respect to τ gives

$$\frac{dW}{d\tau} = \frac{\partial W}{\partial(L+w^*L^*)}\frac{d(L+w^*L^*)}{d\tau} + \frac{\partial W}{\partial\hat{c}}\frac{d\hat{c}}{d\tau} + \frac{\partial W}{\partial\tau}. \tag{A.22}$$

We can note that $\partial W/\partial(L+w^*L^*) > 0$ and $d(L+w^*L^*)/d\tau = d(w^*L^*)/d\tau$. From Appendix A.5, we know that $\Gamma_2(\kappa;\hat{c}) = 0$ can be rewritten as

$$w^*L^* = \frac{L}{[\sigma/(\sigma-1)]\beta(a,b)-1}, \tag{A.23}$$

with $a = \kappa^{1-\sigma}$, $b = \left[\hat{c}^{k-\sigma+1}\left(\nu-\nu_1\hat{c}^{\frac{k-\sigma+1}{k-\sigma+2}}\right)\right]^{-1}$, and $\beta(a,b)$ given in Eq. (A.18). This establishes

$$\frac{d(w^*L^*)}{d\tau} = \frac{d(w^*L^*)}{d\beta(a,b)}\left[\beta_a'(a,b)\frac{da}{d\kappa}\frac{d\kappa}{d\tau} + \beta_b'(a,b)\frac{db}{d\hat{c}}\frac{d\hat{c}}{d\tau}\right]. \tag{A.24}$$

Accounting for $d(w^*L^*)/d\beta(a,b) < 0$, $\beta_a'(a,b) < 0$, $\beta_b'(a,b) > 0$, $da/d\kappa < 0$, $db/d\hat{c} < 0$, and recollecting from Figure 3 that $d\hat{c}/d\tau < 0$, $d\kappa/d\tau > 0$, we can safely conclude that $d(w^*L^*)/d\tau < 0$, which implies $\partial W/\partial(L+w^*L^*) \times d(L+w^*L^*)/d\tau < 0$.

In a next step, we can note that

$$\frac{\partial W}{\partial \hat{c}} = \frac{W}{\sigma-1}\left[\frac{f_e}{f} - \frac{\hat{c}^k}{\nu-\nu_1\hat{c}}\left(\nu\frac{\sigma-1}{k-\sigma+1} - \nu_1\hat{c}\frac{\sigma-2}{k-\sigma+2}\right)\right]^{-1}$$
$$\times\left\{\frac{k\hat{c}^{k-1}}{\nu-\nu_1\hat{c}}\left[\frac{(\sigma-1)\nu}{k-\sigma+1} - \frac{(\sigma-2)\nu_1\hat{c}}{k-\sigma+2}\right] + \frac{\nu_1\hat{c}^k}{(\nu-\nu_1\hat{c})^2}\left[\frac{(\sigma-1)\nu}{k-\sigma+1} - \frac{(\sigma-2)\nu}{k-\sigma+2}\right]\right\} \quad \text{(A.25)}$$

is positive. In view of $d\hat{c}/d\tau < 0$, we therefore have $\partial W/\partial \hat{c} \times d\hat{c}/d\tau < 0$. Accounting for $\partial W/\partial \tau = 0$, we can finally conclude that $dW/d\tau < 0$.

Differentiating Eq. (16) with respect to f gives

$$\frac{dW}{df} = \frac{\partial W}{\partial(L+\hat{w}^*L^*)}\frac{d(L+w^*L^*)}{df} + \frac{\partial W}{\partial \hat{c}}\frac{d\hat{c}}{df} + \frac{\partial W}{\partial f}, \quad \text{(A.26)}$$

To determine the sign of dW/df, we can first note that $w^*L^* = \kappa L^*/\tau$. Since Figure 3 establishes $d\kappa/df < 0$, we get $\partial W/\partial(L+w^*L^*) \times d(L+w^*L^*)/df < 0$. Noting further that $dW/d\hat{c} > 0$ holds according to Eq. (A.25) and considering $d\hat{c}/df < 0$ (again, from Figure 3), it follows that $\partial W/\partial \hat{c} \times d\hat{c}/df < 0$. Finally, we can compute $\partial W/\partial f > 0$. Taking stock, we can thus conclude that the indirect effect of an increase in f through changes in $\hat{c}$ and κ is negative, whereas the direct effect of an increase in f is positive. It is in general not clear which of these counteracting effects is stronger, implying the the total impact of an increase in f on W can be positive or negative. To substantiate this argument, we have shown in a numerical exercise that for a parameter configuration of $k = \sigma = 2$, $\nu = 0.8$, $\nu_1 = 0.2$, $f_e = 10$, $\tau = 1.5$, $L = 100$, and $L^* = 33$, a marginal increase of f from a low value of 0.1 increases source country welfare, whereas a marginal increase of f from a higher value of 0.2 lowers source country welfare.

In the main text, we argue that the possibility of $dW/df > 0$ is the consequence of a deterioration in the source country's (double) factorial terms of trade. To support this argument, we can note that a (double) factorial terms of trade deterioration is reflected by an increase in κ and can determine the welfare effects that would result in a counterfactual scenario, in which κ does not change. To do so, we combine the offshoring indifference condition $\Gamma_1(\hat{c},\kappa) = 0$ with Eq. (16) to rewrite source country welfare as follows

$$W = \left\{\frac{L+\kappa L^*/\tau}{\sigma}\rho_0^{-1}\right]^{\frac{1}{\sigma}}\frac{\sigma-1}{\sigma}, \quad \text{(A.27)}$$

with

$$\rho_0 \equiv \frac{f\hat{c}^{\sigma-1}}{\nu-\nu_1\hat{c}}\left[\frac{\kappa^{1-\sigma}-1}{(1-\sigma)\ln\kappa} - 1\right]^{-1} = f_e - \frac{\hat{c}^k}{\nu-\nu_1\hat{c}}\left[\frac{(\sigma-1)\nu}{k-\sigma+1} - \frac{(\sigma-2)\nu_1\hat{c}}{k-\sigma+2}\right]f. \quad \text{(A.28)}$$

Thereby, the second equality sign in Eq. (A.28) reflects the offshoring indifference condition.

Applying the implicit function theorem to Eq. (A.28), we obtain

$$\left.\frac{d\hat{c}}{df}\right|_{\Gamma_1(\cdot)=0,\kappa=\text{const.}} = -\frac{AE+B}{f(CE+D)}, \tag{A.29}$$

with

$$A \equiv \frac{\hat{c}^{\sigma-1}}{\nu-\nu_1\hat{c}}, \qquad B \equiv \frac{\hat{c}^k}{\nu-\nu_1\hat{c}}\left[\frac{(\sigma-1)\nu}{k-\sigma+1} - \frac{(\sigma-2)\nu_1\hat{c}}{k-\sigma+2}\right], \tag{A.30}$$

$$C \equiv A\left(\frac{\sigma-1}{\hat{c}} + \frac{\nu_1}{\nu-\nu_1\hat{c}}\right), \qquad D \equiv \frac{kB}{\hat{c}} + \frac{\nu_1\hat{c}^k}{(\nu-\nu_1\hat{c})^2}\left[\frac{(\sigma-1)\nu}{k-\sigma+1} - \frac{(\sigma-2)\nu}{k-\sigma+2}\right], \tag{A.31}$$

and

$$E \equiv \left[\frac{\kappa^{1-\sigma}-1}{(1-\sigma)\ln\kappa} - 1\right]^{-1} \tag{A.32}$$

Totally differentiating ρ_0 with respect to f, keeping κ constant, then gives

$$\left.\frac{d\rho_0}{df}\right|_{\kappa=\text{const.}} = AE - CE\frac{AE+B}{CE+D} = \frac{AD-BC}{CE+D}E \tag{A.33}$$

Accounting for

$$AD - BC = \frac{\hat{c}^{k-1}[(\sigma-1)\nu - (\sigma-2)\nu_1\hat{c}]}{\nu-\nu_1\hat{c}}A > 0, \tag{A.34}$$

we can thus safely conclude that $d\rho_0/df|_{\kappa=\text{const.}} > 0$, implying that W unambiguously decreases in f if κ stays constant. This completes the proof. *QED*

A.7 A model variant without overlap

To remove the overlap from the model we can set the probability of offshoring when making the f investment equal to one. More specifically, we can set $\nu_0 = 1$ and $\nu_1 = 0$.Whereas this modification does not affect Eqs. (1)-(3) and Eq. (5), it changes the probability density function of c, which simplifies to

$$g_c(c) = \begin{cases} -\frac{1}{\ln\kappa}c^{k-1}\left[\left(\frac{1}{\kappa}\right)^k - 1\right] & \text{if} \quad c \le \kappa\hat{c} \\ -\frac{1}{\ln\kappa}c^{k-1}\left[\left(\frac{\hat{c}}{c}\right)^k - 1\right] & \text{if} \quad c \in (\kappa\hat{c}, \hat{c}] \\ kc^{k-1} & \text{if} \quad c > \hat{c} \end{cases}, \tag{A.35}$$

and it alters the indifference condition in Eq. (6), which now reads

$$\sigma f = (1-\hat{z})^{1-\sigma} r^d(1)\left[\frac{\kappa^{1-\sigma}-1}{(1-\sigma)\ln\kappa} - 1\right]. \tag{A.36}$$

Computing aggregate revenues as outlined above, we obtain

$$R = M \int_0^1 r(c) g_c(c) dc = M r(1) \frac{k}{k-\sigma+1} \left[1 + \hat{c}^{k-\sigma+1} \left(\frac{\kappa^{1-\sigma} - 1}{(1-\sigma)\ln\kappa} - 1 \right) \right], \tag{A.37}$$

and using the latter in the free entry condition, we obtain the modified *offshoring indifference condition*:

$$\Gamma_1(\hat{c}, \kappa) \equiv \hat{c}^{\sigma-1} \frac{k}{k-\sigma+1} + \left[\hat{c}^k \frac{\sigma-1}{k-\sigma+1} - \frac{f_e}{f} \right] \left[\frac{\kappa^{1-\sigma} - 1}{(1-\sigma)\ln\kappa} - 1 \right] = 0, \tag{A.38}$$

with $d\hat{c}/d\kappa|_{\Gamma_1(\cdot)=0} < 0$.

To get a second link between $\hat{c}$ and κ we can make use of Eq. (8) and compute a modified *labor market condition.* Following the derivation steps from the main text, we obtain

$$\Gamma_2(\kappa, \hat{c}) \equiv \kappa \left\{ \frac{\sigma}{\sigma-1} \frac{1 + \hat{c}^{k-\sigma+1} \left(\frac{\kappa^{1-\sigma}-1}{(1-\sigma)\ln\kappa} - 1 \right)}{\hat{c}^{k-\sigma+1} \frac{1+\kappa^{1-\sigma}[(1-\sigma)\ln\kappa - 1]}{[(1-\sigma)\ln\kappa]^2}} - 1 \right\} - \frac{\tau L}{L^*} = 0. \tag{A.39}$$

with $d\kappa/d\hat{c}|_{\Gamma_2(\cdot)=0} > 0$. For sufficiently high levels of τ and f, the two conditions $\Gamma_1(\hat{c}, \kappa) = 0$ and $\Gamma_2(\kappa, \hat{c}) = 0$ characterize a unique interior equilibrium whose properties are similar to those of the benchmark model.

To complete the characterization of the model variant without overlap, we can further compute

$$r(1) = \sigma f \left[\frac{f_e}{f} - \hat{c}^k \frac{\sigma-1}{k-\sigma+1} \right] \frac{k-\sigma+1}{k} \tag{A.40}$$

and

$$W = \left\{ \frac{L + \kappa L^*/\tau}{\sigma f} \left[\frac{f_e}{f} - \hat{c}^k \frac{\sigma-1}{k-\sigma+1} \right]^{-1} \frac{k}{k-\sigma+1} \right\}^{\frac{1}{\sigma-1}} \frac{\sigma-1}{\sigma}. \tag{A.41}$$

With these insights at hand, we can finally solve for $\chi = \hat{c}^k$,

$$\frac{R^d}{R} = \frac{1 - \hat{c}^{k-\sigma+1}}{1 + \hat{c}^{k-\sigma+1} \left(\frac{\kappa^{1-\sigma}-1}{(1-\sigma)\ln\kappa} - 1 \right)}, \tag{A.42}$$

$$\mu_1(k) = \frac{k}{k+2} \frac{1 - \hat{c}^{k+2}}{1 - \hat{c}^k} - \left(\frac{k}{k+1} \frac{1 - \hat{c}^{k+1}}{1 - \hat{c}^k} \right)^2, \tag{A.43}$$

and

$$\Delta W = 100 \left\{ \left(1 + \frac{\kappa L^*}{\tau L}\right)^{\frac{1}{\sigma-1}} \left[1 - \hat{c}^k \frac{(\sigma-1)}{k-\sigma+1} \frac{f}{f_e}\right]^{\frac{1}{1-\sigma}} - 1 \right\}. \tag{A.44}$$

This completes the proof. *QED*

B Data appendix

B.1 Further descriptives

Table B.1: Employment and offshoring

Decile	No	Yes
1	93.90	6.10
2	88.90	11.10
3	88.86	11.14
4	88.17	11.83
5	86.06	13.94
6	83.19	16.81
7	75.81	24.19
8	55.35	44.65
9	16.34	83.66
Total	81.22	18.78

Table B.2: Nr. of tasks and offshoring

Decile	No	Yes
1	91.75	8.25
2	92.31	7.69
3	92.44	7.56
4	88.86	11.14
5	85.22	14.78
6	80.39	19.61
7	82.90	17.10
8	76.85	23.15
9	64.62	35.38
Total	83.93	16.07

Source: IAB Establishment Panel, covering 20,334 establishment observations for Germany in the years 1998, 2000, 2002; Descriptive statistics refer to own computations, using inverse probability weights.

Table B.3: Manufacturing

Decile	No	Yes
1	80.36	19.64
2	74.79	25.21
3	72.69	27.31
4	62.21	37.79
5	44.20	55.80
6	40.07	59.93
7	25.94	74.06
8	25.61	74.39
9	22.88	77.12
Average	49.86	50.14

Table B.4: Services

Decile	No	Yes
1	86.41	13.59
2	84.86	15.14
3	76.01	23.99
4	67.74	32.26
5	63.10	36.90
6	63.86	36.14
7	62.44	37.56
8	54.09	45.91
9	46.74	53.26
Average	67.25	32.75

Source: IAB Establishment Panel, covering 20,334 establishment observations for Germany in the years 1998, 2000, 2002; Descriptive statistics refer to own computations, using inverse probability weights.

B.2 The definition of tasks

The 2006 BIBB/BAuA Employment Survey reports different workplace activities (Tätigkeiten) at the worker lever in addition to common occupation codes. We make use of 28 different activities which are summarized in Table B.5:

Table B.5: Task definitions

Nr.	Description	possible answers
1	Manufacture, Produce Goods	often/sometimes/never
2	Measure, Inspect, Control Quality	o/s/n
3	Oversee, Control Machinery and Techn. Processes	o/s/n
4	Repair, Maintain	o/s/n
5	Purchase, Procure, Sell	o/s/n
6	Transport, Store, Dispatch	o/s/n
7	Advertise, Promote, Conduct Marketing and PR	o/s/n
8	Organize, Plan, Prepare (others' work)	o/s/n
9	Develop, Research, Construct	o/s/n
10	Train, Teach, Instruct, Educate	o/s/n
11	Gather Information, Investigate, Document	o/s/n
12	Consult and Inform Colleagues within Plant	yes/no
13	Consult and Inform External Clients	y/n
14	Consult and Inform Others	y/n
15	Entertain, Accommodate, Supply of Food	o/s/n
16	Nurse, Look After, Cure	o/s/n
17	Protect, Secure, Guard, Monitor, Regulate Traffic	o/s/n
18	Work with Computer	o/s/n
19	Clean, Eliminate Waste, Recycle	o/s/n
20	Write Computer Programs or Use Macros	y/n
21	Develop Software, Write Software Program, Systems Analysis	y/n
22	Develop or Produce IT-Technology or Hardware	y/n
23	IT Administration of Networks, IT-Systems, Data Banks, Web server	y/n
24	IT Sales and Distribution	y/n
25	Other Tasks related to IT	y/n
26	Advise, Coach or Train Colleagues within plant	y/n
27	Advise, Coach or Train External Clients	y/n
28	Advise, Coach or Train Others	y/n

As a robustness check (see below), we use the task definitions of Becker and Muendler (2015), who distinguish 15 workplace activities that are summarized in Table B.6:

Table B.6: Task definitions according to Becker and Muendler (2015)

Nr.	Description	possible answers
1	Manufacture, Produce Goods	often/sometimes/never
2	Repair, Maintain	o/s/n
3	Entertain, Accommodate, Prepare Foods	o/s/n
4	Transport, Store, Dispatch	o/s/n
5	Measure, Inspect, Control Quality	o/s/n
6	Gather Information, Develop, Research, Construct	o/s/n
7	Purchase, Procure, Sell	o/s/n
8	Program a Computer	yes/no
9	Apply Legal Knowledge	no/basic/expert knowledge
10	Consult and Inform	o/s/n
11	Train, Teach, Instruct, Educate	o/s/n
12	Nurse, Look After, Cure	yes/no
13	Advertise, Promote, Conduct Marketing and PR	y/n
14	Organize, Plan, Prepare (others' work)	y/n
15	Oversee, Control Machinery and Techn. Processes	o/s/n

B.3 Background material for Section 5

To construct an offshoring measure for the German manufacturing industry, we rely on bilateral trade data from he OECD STAN database and aggregate intermediate and capital goods. To this measure, we add 50 percent of mixed end-use and miscellaneous imports (even though this categories are fairly small). To construct a measure for service offshoring, we use information on trade in services from the WITS database of the Worldbank. Although it is the ambition of the Worldbank to provide detailed data on international trade in services in a systematic way for a large country sample over a long time span, the database in its present form has at least two disadvantages for measuring service trade at a disaggregated level: The number of reported categories changes over time, and a significant fraction of services (more than 50 percent in 2008) cannot be attributed to any of the available categories. To deal with these two problems, we use data on service imports for Germany in 2008, which covers the highest number of categories, and construct for this year the share of *service offshoring* in the overall amount of service imports. We then multiply the thus computed share with observed service imports to compute a measure of service offshoring for each year of the period 1990-2014, presuming that the share of service offshoring stayed constant over this period.

Table B.7 provides details on how we construct service offshoring in 2008. Starting point are the three broad (upper-tier) categories *Transportation*, *Travel*, and *Other Services*. For these categories, we determine the share of service imports associated with offshoring. In the case of transportation, these are all services not associated with the transportation of passengers.

Since the disaggregated data does not add up to the value reported at the aggregate level, we introduce symmetric scaling factors to make the data reported at the disaggregated level consistent with the data reported at the aggregate level. Traveling is not related to the idea of offshoring in our model, and we therefore set its offshoring content equal to zero. Finally, we subtract life insurance and pension funding, personal cultural, and recreational services, as well as government services to obtain a measure for service offshoring in the last category. Again, we use scaling factors to make the disaggregated data consistent with the data at the aggregate. Adding up over all categories, we associate 63.69 percent of the service imports in 2008 with offshoring. We assume that this factor stays constant over time and multiply yearly information on total service imports with 0.6369 to compute time-varying values of service offshoring.

Table B.7: Constructing a measure for service offshoring (BSI)

Title	Category	Service Imports	Service Offshoring	Share
Transportation	205	64,998.699	51,895.628	0.798
Travel	236	91,691.797	0.000	0.000
Other services	981	200,653.203	175,696.092	0,876
Total	200	357,343.688	227,591.719	0.637

Source: Worldbank International Trade in Services Database. Service imports in million USD for Germany in 2008.

Tables B.8 and B.9 display the parameter estimates (Columns 2-4), the observed changes in offshoring openness (Column 5), the intensive margin of the changes in offshoring openness (Column 6) and the welfare effects of offshoring (Column 7) for the model variant with and without overlap, respectively. To determine the parameters in Columns 2-4, we use the estimates from the main text, except for those of κ and $\hat{c}$. We then equate the measure of offshoring openness from theory, $w^*L^*/L = \kappa L^*/(\tau L)$, with offshoring openness in the data to compute κ as a function of $\tau L/L^*$. Making use of the two implicit functions $\Gamma_1(\hat{c}, \kappa) = 0$ and $\Gamma_2(\kappa, \hat{c}) = 0$ then allows us to compute the parameter values reported in Columns 2-4.[27] To determine the intensive margin of offshoring openness, we can first use Eq. (8) together with $L + w^*L^* = R$ to express the domestic wage bill as a function of aggregate revenues R:

$$L = R\left\{1 - \frac{\sigma-1}{\sigma}\frac{\hat{c}^{k-\sigma+1}\left(\nu - \nu_1\hat{c}\frac{k-\sigma+1}{k-\sigma+2}\right)\frac{1+\kappa^{1-\sigma}[(1-\sigma)\ln\kappa-1]}{[(1-\sigma)\ln\kappa]^2}}{1+\hat{c}^{k-\sigma+1}\left(\nu - \nu_1\hat{c}\frac{k-\sigma+1}{k-\sigma+2}\right)\left(\frac{\kappa^{1-\sigma}-1}{(1-\sigma)\ln\kappa} - 1\right)}\right\}. \tag{A.45}$$

[27]Fitting the model to observed offshoring openness gives higher values for κ and lower values for $\hat{c}$ than in the baseline estimation.

Combining this with w^*L^* from Eq. (8), we can compute

$$\frac{w^*L^*}{L} = \left[\frac{\sigma}{\sigma-1} \frac{1+\hat{c}^{k-\sigma+1}\left(\nu - \nu_1 \hat{c}\frac{k-\sigma+1}{k-\sigma+2}\right)\left(\frac{\kappa^{1-\sigma}-1}{(1-\sigma)\ln\kappa} - 1\right)}{\hat{c}^{k-\sigma+1}\left(\nu - \nu_1 \hat{c}\frac{k-\sigma+1}{k-\sigma+2}\right)\frac{1+\kappa^{1-\sigma}[(1-\sigma)\ln\kappa - 1]}{[(1-\sigma)\ln\kappa]^2}} - 1 \right]^{-1}. \tag{A.46}$$

for the model variant with overlap. The respective expression for the model variant without overlap can be obtained by setting $\nu_0 = 1$ and $\nu_1 = 0$ in Eq. (A.46). Evaluating the right-hand side at κ_t, $(\tau L/L^*)_t$, $t \in \{1990, ..., 2014\}$, and $\hat{c}_{1990}$ gives offshoring openness for the hypothetical case that the mass of offshoring firms stays constant at its value of 1990.[28] The relative changes of this measure are reported in Column 6. Finally, the welfare effects attributed to the observed changes in offshoring openness are computed, according to Eq. (20) and Eq. (A.44), respectively.

B.4 Additional robustness checks

In this subsection, we check the robustness of our results regarding an expansion of the set of targeted moment conditions and changes in the definition of tasks used for the empirical exercise.

B.4.1 Additional moment conditions

From inspection of Figures 4 and 5, one may conclude that there is scope for further improving the aptitude of our model to capture the overlap and marginal costs in our data, by accounting for additional moment conditions as targets in the MM estimation. In Panel A of Table B.10 we present the estimation results from an MM estimation that targets the same moments as in the baseline specification and, in addition, the share of offshoring firms at the second decile of the revenue distribution. We see that with an exception of ν_1 the parameter estimates remain more or less unaffected by this modification. The overall welfare gain from access to offshoring is slightly increased to 17.11 percent in this case. Also the computed welfare stimulus from offshoring between 1990 and 2014, and the contribution of the intensive margin to the observed increase of offshoring over this period are somewhat higher than in the baseline model and considerably larger than in the model variant without overlap, for which the moment conditions do not change.

In Panel B of Table B.10 we report the parameter values from an MM estimation that additionally accounts for the mean of marginal costs in the set of targeted moment conditions. Adding the new moment condition changes the parameter estimates in both model variants and this has consequences for the computed welfare effects of offshoring, which are more pronounced. However, the insight that ignoring the overlap leads to a downward bias in the welfare effects of offshoring remains unaffected. Furthermore, our previous insights that the welfare gain attributable to the observed increase in German offshoring openness is downward biased, whereas

[28] Lacking firm-level information on incumbent offshoring producers, the definition of the intensive margin in the empirical exercise differs slightly from the respective definition in the theory section.

Table B.8: Parameter estimates, changes in offshoring openness, and welfare effects of offshoring for the period 1990-2014 (with overlap)

Year	κ	c	$\tau L/L^*$	openness	int. margin	welfare effects
1990	0.92	0.84	5.13	0.00	0.00	0.00
1991	0.92	0.84	5.13	-0.07	-0.04	0.00
1992	0.94	0.82	5.59	-7.15	-3.61	-0.44
1993	0.95	0.80	6.21	-15.32	-7.35	-0.90
1994	0.94	0.81	5.90	-11.39	-5.61	-0.68
1995	0.94	0.82	5.57	-6.84	-3.46	-0.42
1996	0.93	0.83	5.49	-5.72	-2.91	-0.35
1997	0.91	0.86	4.78	6.03	3.25	0.38
1998	0.91	0.87	4.55	10.55	5.80	0.68
1999	0.90	0.88	4.38	14.17	7.91	0.92
2000	0.86	0.92	3.49	37.37	22.49	2.57
2001	0.87	0.91	3.64	32.73	19.44	2.23
2002	0.88	0.90	3.77	29.05	17.07	1.96
2003	0.87	0.91	3.63	32.99	19.61	2.25
2004	0.86	0.92	3.44	38.97	23.55	2.69
2005	0.84	0.94	3.07	52.14	32.52	3.67
2006	0.83	0.95	2.81	62.98	40.16	4.51
2007	0.83	0.95	2.80	63.47	40.51	4.54
2008	0.81	0.96	2.60	73.25	47.57	5.31
2009	0.85	0.93	3.26	45.00	27.61	3.13
2010	0.83	0.95	2.84	61.80	39.31	4.41
2011	0.82	0.96	2.66	70.32	45.44	5.08
2012	0.81	0.96	2.63	71.75	46.48	5.19
2013	0.82	0.95	2.68	69.23	44.65	4.99
2014	0.82	0.95	2.71	67.85	43.65	4.88
Average	0.87	0.90	3.95	31.73	20.16	2.26

the contribution of the extensive margin to the observed expansion of offshoring is exaggerated when ignoring the overlap in the data remain unaffected. When including the additional moment condition, our model produces interior solutions with $\kappa, \hat{c} \in (0,1)$ only for the shorter time period of 1990-2009.

Taking stock, we can conclude from the two robustness checks in this subsection that adding moment conditions for improving the fit of our model with the share of offshoring establishments (at various decile positions) or the mean of marginal costs affects the parameter estimates in a non-trivial way. However, the main finding that ignoring the overlap in the data leads to a

Table B.9: Parameter estimates, changes in offshoring openness, and welfare effects of offshoring for the period 1990-2014 (without overlap)

Year	κ	c	$\tau L/L^*$	openness	int. margin	welfare effects
1990	0.9986	0.74	5.54	0.00	0.00	0.00
1991	0.9986	0.74	5.54	-0.07	0.00	0.00
1992	0.9987	0.73	5.97	-7.15	-0.05	-0.20
1993	0.9989	0.71	6.54	-15.32	-0.10	-0.42
1994	0.9988	0.72	6.25	-11.39	-0.07	-0.31
1995	0.9987	0.73	5.95	-6.84	-0.04	-0.19
1996	0.9987	0.73	5.88	-5.72	-0.04	-0.16
1997	0.9984	0.75	5.22	6.03	0.04	0.17
1998	0.9984	0.76	5.01	10.55	0.07	0.29
1999	0.9983	0.77	4.85	14.17	0.09	0.39
2000	0.9978	0.80	4.03	37.37	0.25	1.01
2001	0.9979	0.80	4.17	32.73	0.21	0.88
2002	0.9980	0.79	4.29	29.05	0.19	0.79
2003	0.9979	0.80	4.16	32.99	0.22	0.89
2004	0.9978	0.80	3.98	38.97	0.26	1.05
2005	0.9975	0.82	3.64	52.14	0.35	1.39
2006	0.9973	0.83	3.39	62.98	0.42	1.67
2007	0.9973	0.83	3.38	63.47	0.42	1.69
2008	0.9971	0.84	3.19	73.25	0.49	1.94
2009	0.9976	0.81	3.82	45.00	0.30	1.21
2010	0.9973	0.83	3.42	61.80	0.41	1.64
2011	0.9971	0.84	3.25	70.32	0.47	1.86
2012	0.9971	0.84	3.22	71.75	0.48	1.90
2013	0.9971	0.84	3.27	69.23	0.46	1.83
2014	0.9972	0.84	3.30	67.85	0.45	1.80
Average	0.9979	0.79	4.45	31.73	0.21	0.84

downward bias in the welfare effects of offshoring and to an exaggeration of the extensive margin in explaining the evolution of offshoring is robust to the considered changes in the set of targeted moment conditions.

B.4.2 Alternative task definitions

Whereas the BIBB BAuA database allows to distinguish 28 different tasks conducted by the German workforce, Becker and Muendler (2015) restrict their attention to 15 (somewhat more broadly defined) tasks, for which they have survey information over a longer time horizon. To

Table B.10: Estimation results when accounting for additional data moments

Panel A: Additional moment – share of offshoring firms at the second decile					
	ν_0	ν_1	σ	k	κ
Estimates	0.20 (0.00)	0.22 (0.02)	6.20 (0.03)	5.86 (0.04)	0.70 (0.01)
	ΔW		ΔW 1990-2014		Int. margin 1990-2014
Effects	17.11%		5.22%		80.53%

Panel B: Additional moment – mean of marginal costs (with overlap)					
	ν_0	ν_1	σ	k	κ
Estimates	0.15 (0.01)	0.23 (0.02)	3.17 (0.04)	2.63 (0.03)	0.40 (0.00)
	ΔW		ΔW 1990-2009		Int. margin 1990-2009
Effects	44.01%		10.45%		84.95%

Panel C: Additional moment – mean of marginal costs (without overlap)					
	$\hat{c}$	σ	k	κ	
Estimates	0.30 (0.00)	1.77 (0.00)	1.25 (0.00)	0.98 (0.00)	
	ΔW		ΔW 1990-2009		Int. margin 1990-2009
Effects	19.26%		9.82%		1.20%

Note: Estimates refer to the mean value of 200 bootstraps. Standard errors in parentheses to the right of the estimates.

see whether and to what extent the composition of tasks affects our results, we have redone the MM estimation, using the same task definitions as in Becker and Muendler (2015). We report the results from this exercise in Table B.11, where Panels A and B refer to the model variants with and without overlap, respectively. For the model variant with overlap (Panel A) we find that using the alternative task definitions lowers the parameter estimates in comparison to the baseline specification. Since a lower value of κ is associated with a larger cost saving from offshoring, it is intuitive that the welfare gains due to access to offshoring increase, when relying on the task definitions of Becker and Muendler (2015). Also the welfare gain of Germany due to offshoring over the period 1990-2014 as well as a the contribution of the intensive margin to the overall increase of German offshoring exposure over this period are higher than in the baseline specification.

The parameter estimates for the model variant without offshoring are summarized in Panel B. Again, we find that the alternative definition of tasks leads to lower parameter estimates. The only exception is κ, whose value remains unaffected. The welfare gain from access to offshoring is more pronounced than in the baseline specification and 68.35 percent lower than

Table B.11: Task definitions as in Becker and Muendler (2015)

Panel A: With overlap					
	ν_0	ν_1	σ	k	κ
Estimates	0.10 (0.00)	0.62 (0.01)	3.62 (0.01)	3.86 (0.01)	0.58 (0.00)
	ΔW		ΔW 1990-2014		Int. margin 1990-2014
Effects	27.93%		12.34%		79.49%

Panel B: Without overlap					
	$\hat{c}$	σ	k	κ	
Estimates	0.72 (0.00)	3.79 (0.00)	4.51 (0.01)	0.997 (0.00)	
	ΔW		ΔW 1990-2014		Int. margin 1990-2014
Effects	8.84%		3.64%		0.34%

Note: Estimates refer to the mean value of 200 bootstraps. Standard errors in parentheses to the right of the estimates.

in the model variant with offshoring. The welfare gain from the observed increase in German offshoring openness over the period 1990-2014 is higher, whereas the contribution of the intensive margin to the increase in offshoring openness over this period is slightly lower than in the baseline specification. Both of these values are significantly smaller than for the model variant with overlap, lending support to our finding that ignoring the overlap in the data leads to a downward bias in the welfare effects and an exaggeration of the extensive margin of offshoring. We can therefore conclude that the two main insights from our analysis remain unaffected by the considered change in the definition of tasks.