

Schmerer, Hans-Jörg; Capuano, Stella; Egger, Hartmut; Koch, Michael

**Conference Paper**

## Offshoring and Firm Overlap

Beiträge zur Jahrestagung des Vereins für Socialpolitik 2015: Ökonomische Entwicklung -  
Theorie und Politik - Session: Offshoring I, No. B02-V3

**Provided in Cooperation with:**

Verein für Socialpolitik / German Economic Association

*Suggested Citation:* Schmerer, Hans-Jörg; Capuano, Stella; Egger, Hartmut; Koch, Michael (2015) :  
Offshoring and Firm Overlap, Beiträge zur Jahrestagung des Vereins für Socialpolitik 2015:  
Ökonomische Entwicklung - Theorie und Politik - Session: Offshoring I, No. B02-V3, ZBW - Deutsche  
Zentralbibliothek für Wirtschaftswissenschaften, Leibniz-Informationszentrum Wirtschaft

This Version is available at:

<https://hdl.handle.net/10419/113147>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen  
Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle  
Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich  
machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen  
(insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten,  
gelten abweichend von diesen Nutzungsbedingungen die in der dort  
genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal  
and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to  
exhibit the documents publicly, to make them publicly available on the  
internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content  
Licence (especially Creative Commons Licences), you may exercise  
further usage rights as specified in the indicated licence.*

# Offshoring and Firm Overlap\*

Stella Capuano

IAB Nuremberg

Hartmut Egger<sup>†</sup>

University of Bayreuth  
CESifo, GEP, and IfW

Michael Koch

University of Bayreuth

Hans-Jörg Schmerer

University of Duisburg  
CESifo and IAB

February 7, 2015

## Abstract

We set up a model of offshoring with heterogeneous producers that captures two empirical regularities on offshoring firms: larger, more productive firms are more likely to make use of the offshoring opportunity; the fraction of firms that engages in offshoring is positive and smaller than one in any size or revenue category. These patterns generate an overlap of offshoring and non-offshoring firms, which is non-monotonic in the costs of offshoring. In an empirical exercise, we employ firm-level data from Germany to estimate key parameters of the model. We show that ignoring the overlap leads to a severe downward bias in the estimated gains from offshoring, which amounts to almost 60 percent in our model.

JEL-Classification: F12, F14

Keywords: Offshoring, Heterogeneous firms, Firm overlap

---

\*We are grateful to participants of research seminars at the University of Bergen, the University of Duesseldorf, the University of Heidelberg, and the University of Tuebingen for helpful comments and suggestions.

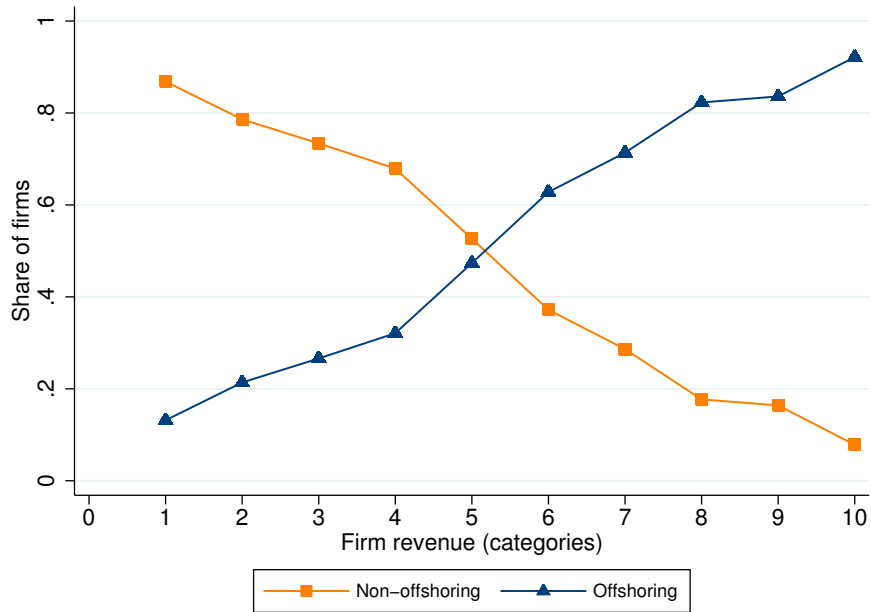
<sup>†</sup>Corresponding author: University of Bayreuth, Department of Law and Economics, Universitätsstr. 30, 95447 Bayreuth, Germany; Email: hartmut.egger@uni-bayreuth.de.

# 1 Introduction

Offshoring and its effects on domestic labor markets have played a prominent role in academic research and public debate over the last two decades. In recent years, attention in the science community has shifted towards understanding the specific nature of firms that engage in production shifting to low-wage countries. Relying on models of heterogeneous firms, trade economists have pointed out that firms which make use of offshoring opportunities systematically differ from their non-offshoring counterparts. They are larger, more productive, and make higher profits (see Antràs and Helpman, 2004; Antràs et al., 2006; Egger et al., 2013). This pattern finds strong empirical support (cf. Hummels et al., 2014; Moser et al., 2014). However, by relying on models with sharp selection into offshoring existing theoretical work misses a considerable overlap of offshoring and non-offshoring firms in the data, which materializes because some but not all firms from any size and revenue category make use of production shifting to low-wage countries.

The purpose of this paper is twofold. In a first step, we construct a theoretical model that allows us to capture the overlap of offshoring and non-offshoring firms. This overlap is documented for German manufacturing firms in Figure 1. The figure confirms that offshoring is conducted by firms from all revenue categories, with the share of offshoring firms increasing in revenues. Tables 1 and 2 show that this finding is independent of the specific performance measure used for categorizing firms. Ranking firms according to their (domestic) employment size or the number of tasks performed at home shows a similar overlap of offshoring and non-offshoring firms, and a positive correlation between firm performance and the likelihood of offshoring in our data-set. In a second step, we estimate key parameters of our model, using firm-level data from German manufacturing industries. Based on these parameter estimates, we then quantify how large the bias in the welfare effects of offshoring is, when ignoring the overlap in the data.

In the theoretical part of the paper, we set up a two-country model of offshoring, with labor being the only factor of production. The two countries differ in their level of development and, to keep things simple, we assume that all firms are headquartered in the more advanced economy, rendering the less developed country a labor reservoir for foreign production (cf. Egger et al., 2013). We can thus associate the more and the less developed economy with the source and the host country of offshoring, respectively. Following Acemoglu and Autor (2011), we model



Source: IAB establishment panel; own computations

Figure 1: Share of offshoring firms in different revenue categories

Table 1: Firm size and offshoring

Employment	No	Yes
1-5	82.21	17.69
6-10	75.43	24.57
11-18	73.84	26.16
19-30	62.47	37.53
31-54	47.12	52.88
55-97	36.56	63.44
98-178	26.31	73.69
179-306	17.03	82.97
307-680	16.10	83.90
> 680	6.76	93.24
Total	45.93	54.07

Table 2: Nr. of tasks and offshoring

Nr. of tasks	No	Yes
1-9	82.91	17.09
10-12	76.65	23.35
13-14	68.00	32.00
15-16	56.86	43.14
17	52.36	47.64
18	30.77	69.23
19-22	45.44	54.56
23	24.92	75.08
24	16.69	83.31
> 24	11.58	88.42
Total	69.29	30.71

Data source: IAB establishment panel, manufacturing firms only. Categories refer to deciles in the size distribution of firms. Descriptive statistics computed using inverse probability weights. Sample size: 8173 firms.

production as the assembly of tasks. Firms differ in the number of tasks performed in the production process. The more tasks a firm performs, the stronger is the division of labor in the manufacturing of goods, with positive implications for labor productivity. Firms also differ in the share of tasks that can be offshored to the low-wage host country.<sup>1</sup> In the tradition of theoretical work building on the Melitz (2003) framework, we model firm heterogeneity as the outcome of a lottery. However, in our framework firms draw *two* technology parameters: the number of tasks and the share of offshorable tasks.

After the lottery, firms are aware of the number of tasks they need to perform in their production process. However, they are not informed about the offshorability of their tasks, and hence they have to form expectations based on the underlying distribution. Conditional on these expectations, firms then decide on making a fixed cost investment in order to explore their offshoring opportunities, and if at least some of their tasks can actually be shifted abroad they will become an offshoring firm with their imports being subject to iceberg transport costs. Firms that are unlucky and end up having not a single task, whose production can be moved abroad, are confined to domestic production, despite their fixed cost investment. Provided that the fixed costs of offshoring are sufficiently high, only the most productive firms will make the investment, and hence our model features self-selection of the most productive firms into offshoring along with an overlap of offshoring and non-offshoring firms for a wide range of revenue levels. This renders the model consistent with the evidence reported in Figure 1, and it still makes the framework flexible enough to account for firm-level responses to changes in the costs of offshoring along two empirically relevant margins: the *extensive margin*, which refers to changes in the composition of offshoring firms; and the *intensive margin*, which refers to changes in the volume of imported goods for a given firm composition.<sup>2</sup>

We use this model to show how changes in variable and fixed offshoring costs affect offshoring behavior and welfare in the source country. A decline in these costs leads to an expansion of offshoring along the extensive and intensive margin. This enhances production possibilities and therefore generates welfare gains for the source country.<sup>3</sup> Being interested in the offshoring

---

<sup>1</sup>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.

<sup>2</sup>The relevance of the extensive and the intensive margins of trade are well documented by Hummels and Klenow (2005). Bergin et al. (2011) provide evidence for the significance of the often ignored expansion of offshoring at the extensive margin.

<sup>3</sup>There are also welfare gains in the host country, because labor demand there increases with lower offshoring

overlap, we construct an index to measure this overlap. This index has a minimum value of zero if either no or all firms conduct offshoring within a certain (marginal) cost category, and it reaches a maximum value of one if 50 percent of the producers in this category are offshoring firms. We aggregate the index to an economy-wide measure of overlap, which is non-monotonic in offshoring costs. It decreases in offshoring costs if only the most-productive firms find it attractive to shift task production abroad, whereas it increases in offshoring costs if production shifting becomes common practice of all firms.

In the second part of the paper, we combine data from the linked employer-employee data-set and the establishment panel of the Institute for Employment Research (IAB) with information from the BIBB-BAuA data-set on task classifications to estimate key parameters of our model. For this purpose, we employ a methods of moments estimation strategy, in which we minimize the distance between selected moments from the data and the model (see Cameron and Trivedi, 2005, for a discussion of this approach). With the respective parameter values at hand, we then assess the model fit, by comparing the overlap of offshoring and non-offshoring firms predicted by our model with the overlap from the data. Our model explains about 35 percent of the overlap in the data, but it underestimates the overlap at the upper and the lower tail of the revenue distribution.

We also analyze to what extent ignoring the overlap in the data matters for our analysis. To tackle this issue, we construct an otherwise identical model without overlap and estimate the parameters of this model without overlap, following the methods of moments approach outlined above. This exercise shows that ignoring the overlap significantly lowers both the costs-saving effects of offshoring and the share of firms making the investment to explore their offshoring opportunities. Furthermore, the differences in the welfare effects of offshoring predicted by the two model variants are huge. Ignoring offshoring in the data leads to a downward bias in the computed gains from offshoring for Germany of almost 60 percent.

In a final step, we conduct a counterfactual analysis to show how changes in offshoring costs affect welfare and overlap. Whereas an increase in the costs of offshoring lowers welfare monotonically, a decomposition analysis shows that the relative importance of adjustments along the intensive and extensive margin changes over the path of offshoring. More specifically, the extensive margin of offshoring is the engine for welfare gains at high levels of offshoring costs and

---

costs, leading to a higher real wage. However, choosing a parsimonious structure of the host country to keep the model analytically tractable, we do not put emphasis on the welfare effects in this economy.

only few offshoring firms, whereas the intensive margin of offshoring becomes the main source of welfare gain for low levels of offshoring costs, when production shifting becomes common practice among all producers. The counterfactual analysis confirms the theoretical insight that the impact of offshoring costs on overlap is non-monotonic. The overlap reaches a maximum value of 24.5 percent which is almost one half of the overlap observed in our data-set.

By emphasizing the overlap of offshoring and non-offshoring firms, our analysis is related to Schröder and Sørensen (2012), who point to an overlap in firm productivity when considering the export activity of firms. Their claim is that the sharp selection of firms into exporting in the Melitz (2003) model is an artifact of ranking firms according to their marginal instead of total factor productivity, which is the productivity measure that is commonly used in empirical work. Once total factor productivity is accounted for, the Melitz model generates an overlap in the productivity distribution of firms located in the neighborhood of the marginal productivity cutoff that separates exporting from non-exporting firms in this framework. Whereas this insight is useful for making the Melitz model with the mixed empirical evidence on productivity premia of exporting firms (cf. Bernard and Jensen, 1995; Greenaway et al., 2005; Chang and van Marrewijk, 2011), it is of no help for explaining the overlap of offshoring and non-offshoring firms, when ranking firms according to their revenues instead of productivities (see Figure 1). Furthermore, extending the Melitz (2003) model to allow firms drawing two correlated parameters in the technology lottery relates our analysis to Davis and Harrigan (2011) and Harrigan and Reshef (2015). These authors use insights from the theory of copulas (cf. Nelson, 2007) to motivate the functional form of the joint distribution of technology parameters. Whereas this approach provides considerable flexibility regarding the specific correlation of the two technology parameters, we follow a different path and model the correlation of technology parameters explicitly to render the model analytically solvable and to get clear empirically testable hypotheses from our model.

The remainder of the paper is organized as follows. In Section 2 we set up the theoretical model and introduce the offshoring measure used in our analysis. In Section 3 we describe the data-set, calibrate key model parameters and discuss the goodness of fit of our model. 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 our data-set. Section 5 presents the counterfactual analysis, which shows how changes in offshoring costs affect welfare and overlap in our model. 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  $q(\omega)$ . The representative consumer's utility is given by

$$U = \left[ \int_{\omega \in \Omega} q(\omega)^{\frac{\sigma-1}{\sigma}} d\omega \right]^{\frac{\sigma}{\sigma-1}}, \quad (1)$$

where  $\sigma > 1$  is the elasticity of substitution between different varieties  $\omega$  and  $\Omega$  is the set of available consumer goods. Maximizing  $U$  subject to the representative consumer's budget constraint  $I = \int_{\omega \in \Omega} p(\omega)q(\omega)$  gives isoelastic demand for variety  $\omega$ :

$$q(\omega) = \frac{I}{P} \left[ \frac{p(\omega)}{P} \right]^{-\sigma}, \quad (2)$$

where  $I$  is aggregate income,  $p(\omega)$  is the price of good  $\omega$  and

$$P = \left[ \int_{\omega \in \Omega} p(\omega)^{1-\sigma} d\omega \right]^{\frac{1}{1-\sigma}} \quad (3)$$

is a 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 economy is the host country of offshoring, whereas the more advanced economy is the source country of offshoring. We adopt the modeling strategy of Egger et al. (2013) and 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. Firms perform different tasks, which are



combined in a Cobb-Douglas production technology to produce output  $q$ :<sup>4</sup>

$$q = \frac{z}{1-z} \exp \frac{1}{z} \int_0^z \ln x(i) di, \quad (4)$$

where  $x(i)$  denotes the output for task  $i$  and  $z$  is the length of the task interval, i.e. the mass of the tasks, performed by the firm. The technology in Eq. (4) captures in a simple way the 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  $C^v = \int_0^z \zeta(i)x(i)di$ , where  $\zeta(i)$  is the effective labor cost, which equals the domestic wage  $w$  if a task is performed at home and 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, i.e.  $\zeta(i)x(i) = \zeta x$  for all  $i$ . Furthermore, the marginal production cost of a firm that performs all tasks at home (superscript  $d$ ) and a firm that offshores a share  $s$  of its tasks to the host country (superscript  $o$ ) are given by

$$c^d = (1-z)w, \quad c^o = (1-z)w\kappa^s, \quad (5)$$

respectively, where  $\kappa \equiv \tau w^*/w$  denotes the *effective* wage differential. 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  $\kappa^s$  with the marginal *cost saving effect* of offshoring (cf. Egger et al., 2013).

To enter the source country, firms must make an initial investment of  $f_e$  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 tuple  $(z, s)$  with  $z$  and  $s$  being the length of the task interval and the share of tasks that are offshorable, respectively. The length of the task interval  $z$  is Pareto distributed over the unit interval with a probability density function  $f_z(z) = k(1-z)^{k-1}$ . The distribution of  $s$  is more sophisticated and 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  $z$ . If a firm has some

---

<sup>4</sup>To save on notation, we use the same variable for production and consumption and suppress firm indices, where this is possible without generating confusion.

offshorable tasks, the share of offshorable tasks  $s$  is uniformly distributed over the interval  $[0, 1]$ . The interdependence of the two random variables  $z$  and  $s$  allows us to capture the two stylized facts documented in Table 2: (i) firms which perform more tasks are more likely to offshore the production of some of these tasks; (ii) for a given length of the task interval only a subset of firms engages in offshoring. After the lottery, firms are informed about their  $z$ -level, but they cannot observe which and how many tasks are offshorable. However, firms form expectations on  $s$ , i.e. on the potential of their technology for and the gains from offshoring. Depending on their  $z$ -level, they can then invest  $f$  units of labor into a fixed offshoring service input that provides information on the share  $s$  of offshorable tasks and the type of tasks that can be moved offshore.

We consider a static model and assume that firms are active, irrespective of their  $z$ -draw. By maximizing profits, firms set prices as a constant markup  $\sigma/(\sigma - 1)$  over their marginal costs,  $c$ :  $p = c\sigma/(\sigma - 1)$ , where  $c = c^d$  if the firm produces purely domestically and  $c = c^o$  if the firm moves a share  $s$  of its tasks abroad. Revenues are given by  $r = IP^{\sigma-1}p^{1-\sigma}$ , and in view of constant markup pricing, we can therefore determine the following two revenue ratios:

$$\frac{r^i(z_1)}{r^i(z_2)} = \left( \frac{1 - z_1}{1 - z_2} \right)^{1-\sigma}, \quad \frac{r^o(z)}{r^d(z)} = \kappa^{s(1-\sigma)}. \quad (6)$$

The first expression gives the revenues ratio of two firms with the same offshoring status  $i \in (d, o)$  but differing  $z$ -length, and it shows that higher revenues are realized by the firm which makes use of more tasks in its production process. The second expression gives the revenue ratio of two firms with the same  $z$ -length but differing offshoring status, and it shows that higher revenues are realized by the offshoring firm due to the cost-saving involved in shifting production abroad. According to Eq. (6), the position of firms in the revenue distribution is fully characterized by their  $z$  and  $s$  draw.

As outlined above, firms must invest  $f$  units of labor to explore their offshoring status, and this is attractive if for a firm with task length  $z$  the expected profit with offshoring  $\mathbb{E}[\pi^o(z, s)] = (1 - z)r^d(z)/\sigma + z\mathbb{E}[r^o(z)/\sigma] - f$  exceeds the profit without offshoring  $\pi^d(z) = r^d(z)/\sigma$ . Hence, firms make the investment if  $z \{ \mathbb{E}[r^o(z) - r^d(z)] \} \geq \sigma f$  or, in view of Eq. (6), if  $z(1 - z)^{1-\sigma}r^d(1) \{ \mathbb{E}[\kappa^{s(1-\sigma)}] - 1 \} \geq \sigma f$ . It is immediate that the attractiveness of offshoring investment increases with task length  $z$ , and this allows to determine a unique  $\hat{z}$  that renders firms indifferent between investing and not investing  $f$ . Accounting for  $\mathbb{E}[\kappa^{s(1-\sigma)}] = \int_0^1 \kappa^{s(1-\sigma)} ds$ ,

the indifferent firm is characterized by the following condition

$$\sigma f = \hat{z}(1 - \hat{z})^{1-\sigma} r^d(1) \left[ \frac{\kappa^{1-\sigma} - 1}{(1 - \sigma) \ln \kappa} - 1 \right]. \quad (7)$$

## 2.2 The general equilibrium

To solve for the general equilibrium, we choose source country labor as numéraire and set the respective wage rate equal to one:  $w = 1$ . Since marginal costs contain all relevant information of a firm's success in the task lottery, we can rank firms according to their  $c$ -level, with the advantage that we do not have to further distinguish between the offshoring and domestic mode of production. Put differently, we can omit indices  $d$  and  $o$  in the subsequent discussion when capturing firm heterogeneity by differences in  $c$ . The marginal cost of the least productive firm – which has task length 0 and is therefore a purely domestic producer – is given by  $c = 1$ . The marginal cost of all producers with  $z < \hat{z}$  is given by  $c = 1 - z$  and hence there is no difference if we rank purely domestic producers according to  $z$  or  $c$ . Things are different however for offshoring firms, which have a task length of  $z \geq \hat{z}$ . For these firms marginal costs are either given by  $c = 1 - z$  if none of their tasks is offshorable or by  $c = (1 - z)\kappa^s$  if a share  $s$  of their tasks can be produced abroad. In the latter case, marginal costs are the product of two random variables, and hence the ranking of  $c$  cannot be directly inferred from the ranking of  $z$ . The picture becomes even more complicated, when taking into account that the probability of offshoring for a firm with task length  $z \geq \hat{z}$  depends on the number of tasks operated by the firm, because this implies that 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 distribution of  $c$ . There, we show that the probability density function (pdf) of marginal production costs  $c$  is given by

$$f_c(c) = \begin{cases} kc^k - \frac{1}{\ln \kappa} \left\{ c^{k-1} \left[ \left( \frac{1}{\kappa} \right)^k - 1 \right] - \frac{kc^k}{k+1} \left[ \left( \frac{1}{\kappa} \right)^{k+1} - 1 \right] \right\} & \text{if } c \leq \kappa \hat{c} \\ kc^k - \frac{1}{\ln \kappa} \left\{ c^{k-1} \left[ \left( \frac{\hat{c}}{c} \right)^k - 1 \right] - \frac{kc^k}{k+1} \left[ \left( \frac{\hat{c}}{c} \right)^{k+1} - 1 \right] \right\} & \text{if } c \in (\kappa \hat{c}, \hat{c}] \\ kc^{k-1} & \text{if } c > \hat{c} \end{cases} \quad (8)$$

with  $\hat{c} \equiv 1 - \hat{z}$ .

Figure 2 displays  $f_c(c)$ . As we can see from this figure,  $f_c(c)$  has support on the unit interval and features a discontinuity at  $\hat{c}$ . This is because at cutoff cost level  $\hat{c}$  the investment into  $f$

becomes attractive, so that a subset of firms with task length  $z \geq \hat{z}$  starts offshoring. These firms experience a cost saving effect, and hence end up with  $c < \hat{c}$ . Put differently, offshoring shifts firms towards lower marginal costs, and this explains the discontinuity in Figure 2. The kink of the pdf function at  $\kappa\hat{c}$  is also a result of offshoring firms being shifted towards lower marginal costs. However, the type of firms making use of offshoring is confined by condition  $z \geq \hat{z}$ . Non-offshoring firms are not shifted in the  $c$ -distribution and this constrains  $f_c(c)$  for all  $c \in (\kappa\hat{c}, \hat{c}]$ .

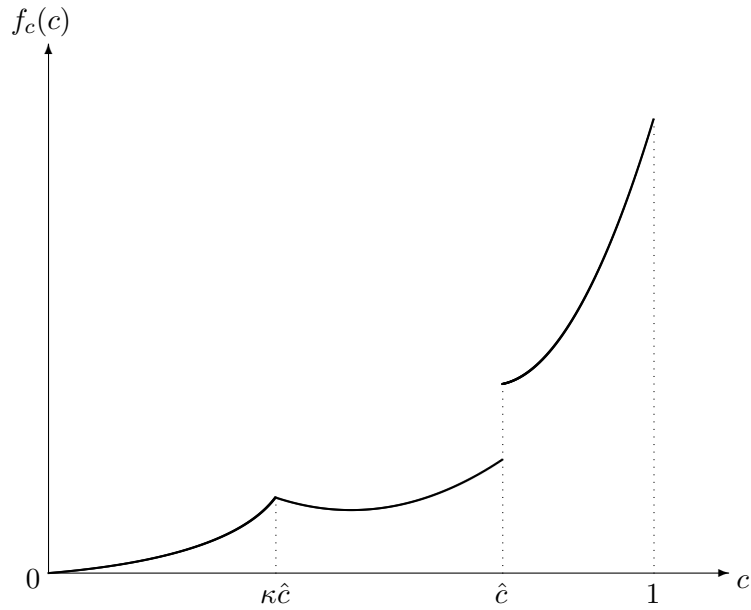


Figure 2: The probability density function  $f_c(c)$

With the distribution of marginal costs at hand, we can compute economy-wide revenues. As formally shown in the appendix, this gives

$$\begin{aligned}
 R &= M \int_0^1 r(c) f_c(c) dc \\
 &= Mr(1) \left[ \frac{k}{k-\sigma+1} + \hat{c}^{k-\sigma+1} \left( \frac{k}{k-\sigma+1} - \hat{c} \frac{k}{k-\sigma+2} \right) \left( \frac{\kappa^{1-\sigma} - 1}{(1-\sigma) \ln \kappa} - 1 \right) \right], \quad (9)
 \end{aligned}$$

where  $k > \sigma - 1$  has been assumed to ensure a finite positive value of  $R$ . Since free entry implies that firms make zero profits on average, constant markup pricing establishes  $R = M\sigma(f_e + \hat{c}^k f)$ . Accounting for Eq. (7), we can solve for  $r(1)$  as a function of  $\hat{c}$  and  $\kappa$ . Substitution into Eq. (9)

gives *offshoring indifference condition* (OC)

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

As formally shown in the appendix,  $\Gamma_1(\cdot) = 0$  establishes a negative link between  $\hat{c}$  and  $\kappa$ . The larger is  $\kappa$ , 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. As outlined above a higher productivity is associated with a higher task length and thus a lower marginal production cost  $c$  in the absence of offshoring. Intuitively, if the cost saving effect of offshoring vanishes due to  $\kappa = 1$ , all firms prefer domestic production and hence  $\hat{c} = 0$ . In contrast, if the cost saving effect of offshoring goes to infinity due to  $\kappa = 0$ ,  $\hat{c}$  reaches a maximum at  $\hat{c}_1 < 1$ , where  $\hat{c}_1$  is implicitly determined by<sup>5</sup>

$$\frac{\hat{c}_1^k}{1 - \hat{c}_1} \left[ \frac{\sigma - 1}{k - \sigma + 1} - \hat{c}_1 \frac{\sigma - 2}{k - \sigma + 2} \right] = \frac{f_e}{f}. \quad (10)$$

To obtain a second link between  $\hat{c}$  and  $\kappa$  we can use the adding up condition for foreign wages. Starting from the observation that a firm's total expenditures for variable labor input,  $C^v$ , are proportional to the firm's revenues  $r$ , with the factor of proportion being given by  $(\sigma - 1)/\sigma$ , we can write  $C^v(c) = [(\sigma - 1)/\sigma]r(c)$ . Due to the Cobb-Douglas technology in Eq. (4), an offshoring firm's expenditure for foreign labor input is given by  $sC^v(c)$ , with  $s$  denoting the share of offshored tasks. Since foreign workers can only be employed in the production of offshored tasks, we can write total labor income in the host country of offshoring as

$$w^*L^* = M\mathbb{E}(s) \frac{\sigma - 1}{\sigma} \int_0^{\hat{c}} r(c) \hat{f}_c(c) dc = M\mathbb{E}(s) \frac{\sigma - 1}{\sigma} r(1) \int_0^{\hat{c}} c^{1-\sigma} \hat{f}_c(c) dc, \quad (11)$$

where  $\mathbb{E}(s) = 1/2$  is the expected value of  $s$  if at least some of a firm's tasks are offshorable and  $\hat{f}_c(c) = f_c(c) - kc^k$  is the density of offshoring producers. Solving the integral and replacing  $r(1)$  from above gives

$$R = w^*L^* \frac{2\sigma}{\sigma - 1} \left\{ 1 + \frac{(1 - \sigma) \ln \kappa}{\kappa^{1-\sigma} - 1} \left[ \frac{k - \sigma + 2}{\hat{c}^{k-\sigma+1} [1 + (1 - \hat{c})(k - \sigma + 1)]} - 1 \right] \right\}. \quad (12)$$

---

<sup>5</sup>One may have expected that  $\hat{c}$  converges to one if  $\kappa$  falls to expected profit gain from offshoring must go to infinity in this case. However, there is a counteracting effect, because  $r(1)$  falls to zero if  $\kappa = 0$  (see below), and with the expected profit gain from offshoring being proportional to  $r(1)$  this counteracting effect leads to  $\hat{c}_1 < 1$ .

Noting further that  $R = L + w^*L^*$  and replacing  $w^*$  by  $\kappa/\tau$  gives the *labor market constraint* (LC):

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

As formally shown in the appendix,  $\Gamma_2(\cdot) = 0$  establishes a positive link between  $\kappa$  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. A larger demand for foreign workers drives up foreign wages and thereby raises  $\kappa$ . If  $\hat{c}$  falls to zero, there is no offshoring, and in this case the host country becomes inactive and  $w^*$  as well as  $\kappa$  fall to zero. In contrast,  $\kappa$  reaches its maximum at a high level of  $\hat{c}$ . Two cases can be distinguished, depending on the ranking of  $\mu \equiv L^*[\sigma + 1 + 2\sigma(k - \sigma + 1)] - \tau L(\sigma - 1) > , = , < 0$ . If  $\mu < 0$ ,  $\kappa$  approaches  $\kappa_2 < 1$  when  $\hat{c}$  goes to one, with  $\kappa_2$  being implicitly given by

$$\kappa_2 \left[ \frac{\sigma + 1}{\sigma - 1} + \frac{2\sigma}{\sigma - 1} \frac{(1 - \sigma) \ln \kappa_2}{\kappa_2^{1-\sigma} - 1} (k - \sigma + 1) \right] = \frac{\tau L}{L^*}. \quad (13)$$

If  $\mu > 0$ ,  $\kappa$  approaches one when  $\hat{c}$  goes to  $\hat{c}_2 < 1$ , with  $\hat{c}_2$  being implicitly given by

$$\frac{\sigma + 1}{\sigma - 1} + \frac{2\sigma}{\sigma - 1} \left[ \frac{k - \sigma + 2}{\hat{c}_2^{k-\sigma+1} [1 + (1 - \hat{c}_1)(k - \sigma + 1)]} - 1 \right] = \frac{\tau L}{L^*}. \quad (14)$$

Irrespective of the specific parameter constellation,  $\Gamma_1(\cdot) = 0$  and  $\Gamma_2(\cdot) = 0$  give a system of two equations which jointly determine a unique interior equilibrium with  $\hat{c}, \kappa \in (0, 1)$ . This equilibrium is depicted by the intersection point of *OC* and *LC* in Figure 3. As illustrated, an increase in variable offshoring costs  $\tau$  rotates the *LC*-locus counter-clockwise, leading to an increase in  $\kappa$  and a decline in  $\hat{c}$ . In contrast, an increase in fixed offshoring cost  $f$  rotates the *OC*-locus clockwise, and this lowers both  $\hat{c}$  and  $\kappa$ .

### 2.3 Welfare and offshoring overlap

Setting  $\hat{z} = 1 - \hat{c}$  in Eq. (7) and combining the resulting expression with  $\Gamma_1(\cdot) = 0$ , we can solve for the revenues of the least productive firm:

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

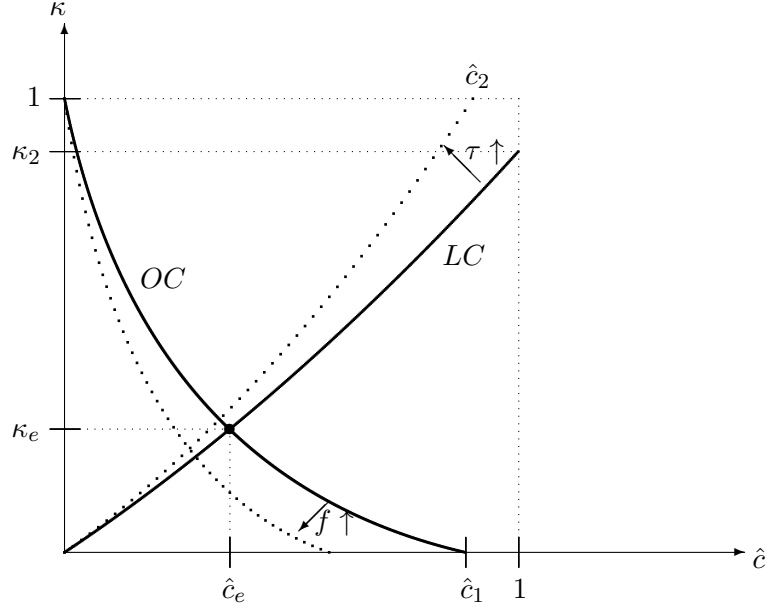


Figure 3: Equilibrium values of  $\hat{c}$  and  $\kappa$

According to Eq. (15), revenues of the marginal firm,  $r(1)$ , decrease in  $\hat{c}$ , and, in view of Eq. (10), they are strictly positive in an interior equilibrium with  $\hat{c}, \kappa \in (0, 1)$ . Source country welfare,  $W$ , is given by per-capita income, which, in view of  $w = 1$ , equals  $P^{-1}$ . To determine the price index, we can start from the observation that  $r(1) = IP^{\sigma-1}[(\sigma-1)/\sigma]^{\sigma-1}$ . Accounting for  $I = L + w^*L^*$  and Eq. (15), we can compute

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

In the appendix, we show that  $W = P^{-1}$ , declines in both  $\tau$  and  $f$ , implying that offshoring provides a welfare stimulus in our setting. The intuition for this result is similar to other settings that feature monopolistic competition between heterogeneous firms. The factor allocation in the one sector economy is efficient, and hence offshoring simply expands the production possibilities with gains for the source country (cf. Dhingra and Morrow, 2013).<sup>6</sup>

To obtain a measure for the overlap of offshoring and non-offshoring firms, we compute for

<sup>6</sup>For completeness, we can also determine the mass of firms. Substituting  $R = L + w^*L^*$  in Eq. (9) and accounting for (15), we can solve for  $M = (L + \kappa L^*/\tau)/[\sigma(f_e + \hat{c}^k f)]$ .

$c \leq \hat{c}$

$$O(c) = 1 - \left| 2 \frac{f_c(c) - kc^k}{f_c(c)} - 1 \right| = 1 - \left| 1 - 2 \frac{kc^k}{f_c(c)} \right|. \quad (17)$$

Substitution of  $f_c(c)$  from Eq. (8), we can compute

$$O(c) = \begin{cases} 1 - \left| \frac{c^{-1}[(1/\kappa)^k - 1] - [k/(k+1)][(1/\kappa)^{k+1} - 1] + k \ln \kappa}{c^{-1}[(1/\kappa)^k - 1] - [k/(k+1)][(1/\kappa)^{k+1} - 1] - k \ln \kappa} \right| & \text{if } c \leq \kappa \hat{c} \\ 1 - \left| \frac{c^{-1}[(\hat{c}/c)^k - 1] - [k/(k+1)][(\hat{c}/c)^{k+1} - 1] + k \ln \kappa}{c^{-1}[(\hat{c}/c)^k - 1] - [k/(k+1)][(\hat{c}/c)^{k+1} - 1] - k \ln \kappa} \right| & \text{if } c \in (\kappa \hat{c}, \hat{c}] \end{cases} \quad (18)$$

As formally shown in the appendix,  $O(c)$  is hump-shaped, reaching a minimum value of zero at  $c = 0$  or  $c = 1$  and maximum value of one at some  $c \in (0, \hat{c})$ . An economy-wide measure of the overlap can then be computed according to  $O = F_c(\hat{c})^{-1} \int_0^{\hat{c}} O(c) f_c(c) dc$ . Considering Eq. (17) and  $f_c(c)$  from Eq. (8), we can compute

$$O = 1 - \frac{1}{F_c(\hat{c})} \int_0^{\kappa \hat{c}} \left| -\frac{1}{\ln \kappa} \left\{ c^{k-1} \left[ \left( \frac{1}{\kappa} \right)^k - 1 \right] - \frac{kc^k}{k+1} \left[ \left( \frac{1}{\kappa} \right)^{k+1} - 1 \right] \right\} - kc^k \right| dc \\ - \frac{1}{F_c(\hat{c})} \int_{\kappa \hat{c}}^{\hat{c}} \left| -\frac{1}{\ln \kappa} \left\{ c^{k-1} \left[ \left( \frac{\hat{c}}{c} \right)^k - 1 \right] - \frac{kc^k}{k+1} \left[ \left( \frac{\hat{c}}{c} \right)^{k+1} - 1 \right] \right\} - kc^k \right| dc. \quad (19)$$

The impact of falling offshoring costs  $\tau$  and  $f$  on the overlap measure  $O$  is non-monotonic. If the respective costs are high, only few firms with high  $z$  make the investment  $f$  and the overlap is small. A decline in offshoring costs allows additional firms to make the  $f$ -investment and since for these newly offshoring firms the overlap is higher than for the incumbent ones, the economy-wide measure of overlap  $O$  increases. Things are different if offshoring costs are already low and hence a significant fraction of firms, including some with low  $z$ , make the  $f$ -investment. A further decline in offshoring costs now adds new offshoring firms for whom the overlap is relatively small and this lowers  $O$ . Of course, there are additional effects at the intensive margin as newly offshoring firms are spread over a whole interval of  $c$ . However, these additional effects do not alter the general picture of a non-monotonic relationship between offshoring costs and offshoring overlap.



### 3 An empirical exercise

To estimate the main parameters of our model, we combine three different data sources: firm information from the IAB establishment panel, worker-level data from the administrative employment records, and occupational task-content from the BIBB-BAuA 2006 data. Firm and worker data can be combined into a matched employer-employee data set, the so-called LIAB. Detailed information about the workers' occupations allows us to construct different task-measures at the firm-level. A detailed description of the data used for the parameter estimation is given in the next subsection.

#### 3.1 Data sources and descriptives

Two variables are particularly relevant for the quantitative exercise: the share of offshoring firms in the sample and an empirical measure for task length  $z$ . The former piece of information was collected in three years of the IAB establishment panel survey: 1999, 2001 and 2003. In the survey, employers were asked whether and where they purchased inputs from external sources in the previous business year. Similar to Moser et al. (2014), we define offshoring as the purchase of intermediates or other inputs from abroad in the previous business year. However, in contrast to Moser et al. (2014) we confine our analysis to manufacturing industries only. Furthermore, we do not exploit the time-series variation and collapse the observations of the three years to a pooled cross section. In total, this results in a data-set with 8,330 manufacturing firms. However, since small firms sometimes report data of low quality, we only keep firms with an employment level of at least five (part- or full-time) workers. Our final sample covers 7,260 firms, with 7,110 of them reporting reliable information on their offshoring status.

As for the length of the task interval, we construct it using the BIBB BIBB-BAuA 2006 "Survey of the Working Population on Qualification and Working Conditions in Germany" (see Rohrbach-Schmidt, 2009). In this survey, individuals are asked about the frequency with which they perform given tasks.<sup>7</sup> For each of the 29 questions regarding the task-content of their job, respondents answer whether they perform that task "often", "sometimes" or "never". Since the unit of analysis in the BIBB data is the individual and not the occupation, it may well be the case that individuals in the same occupation provide slightly different answers about the

---

<sup>7</sup>Several studies have already used the BIBB survey to measure the task-content of occupations, see for example Spitz-Oener (2006) and Becker et al. (2013).

content of their job. Thus, we have to apply some criterion according to which tasks can be straightforwardly assigned to occupations. Specifically, we link an occupation to a given task if at least 60 percent of the interviewees in that occupation declare to perform that task “often” or “sometimes”. In this way we are able to construct a dataset where the observational unit is an occupation, and for each one we can establish how many of the 29 tasks belong to it.

We match the above dataset with the LIAB (IAB linked employer-employee) data, using the occupational classification as the key matching variable. This allows us to compute the number of unique tasks performed in every firm by simply counting the tasks attached to each occupation. For consistency with our theoretical model, we need a measure of the length of the task interval ranging between 0 and 1. We obtain that by simply dividing the number of tasks in each firm by the total number of tasks (i.e. 29). We finally link the thus constructed task information to the establishment data discussed above. Finally, we apply the inverse probability weights of the IAB to make the sample representative for German establishments (cf. Fischer et al., 2008). Table 3 reports the descriptives of firms in our data-set.

Table 3: Summary statistics on firm offshoring status and number of tasks

	Mean	Median	Std. Dev.	Min.	Max.
Offshoring	0.38	0.00	0.49	0	1
Nr. of tasks	13.98	14.00	4.18	1	26
Nr. of tasks/total nr. tasks	0.48	0.48	0.14	0.034	0.90
Nr. of full-time workers	50.88	14.00	294.05	0	48,957
Revenues	9,420,030	1,186,826	98,268,97	10,918.34	12,054,905,371
Value added	3,929,208	612,191	36,265,54	7,387.10	6,718,740,992

IAB establishment panel, years 1999, 2001 and 2003; manufacturing firms with more than four employees. Descriptive statistics computed using inverse probability weights. Value added, revenues and wage bill are in Euro.

The share of offshoring firms in our data-set is significantly higher than the share of offshoring firms reported by Moser et al. (2014). The reason for this difference is that we focus on manufacturing firms only, and offshoring is more common among manufacturing than non-manufacturing firms in the IAB establishment panel. Furthermore, dropping firms with less than five (part- or full-time) employees contributes to a further increase in the share of offshoring firms in our data-set. The variance in both the number of tasks and the number of full-time employees is fairly high, which confirms that firm heterogeneity is an important feature in our data-set. It

is notable, that some firms just perform a single task and have zero full-time workers. This indicates that dropping firms with less than five (part- or full-time) employees still leaves fairly small firms in our data-set. In the last two rows of Table 3, we report revenues and value added. Whereas these measures differ in size, they are strongly positively correlated in our data-set. In the main text, we consider revenues for calibrating key model parameters, as suggested by our model. However, since there are strong vertical linkages between firms in our data-set, which are not captured by our model, we calibrate the model parameters using value added instead of revenues as a robustness check, with the results presented in the appendix.

### 3.2 Estimation strategy

For our estimation we use in particular information on domestic producers, because it is only these firms for which we can directly observe the number of tasks performed in the production process in our data-set. Two key parameters of our model are  $k$  and  $\hat{c}$ , and we can calibrate these parameters by combining the following two equations of our model. The first one is the probability density function of  $c$  for domestic producers

$$f_c^d(c) = \begin{cases} kc^k & \text{if } c \leq \hat{c} \\ kc^{k-1} & \text{if } c > \hat{c} \end{cases} \quad (20)$$

Whereas marginal production costs  $c$  are not directly observable in the data, we can make use of the link  $c = 1 - z$  from the model, and use observed values for the task length in order to infer the required information on  $f_c^d(c)$ . The second equation is the adding up condition for the share of offshoring firms,  $\chi$ .

$$\chi = 1 - \int_0^{\hat{c}} f_c^d(c)dc = 1 - \int_0^{\hat{c}} kc^k dc - \int_{\hat{c}}^1 kc^{k-1} dc = \hat{c}^k \left[ 1 - \frac{k}{k+1} \hat{c} \right]. \quad (21)$$

Based on the data for  $1 - z$  and  $\chi$ , we construct three model moments, i.e. the first and the second moments of the distribution of a firm's length of the task interval and the share of offshoring firms  $\chi$ . In order to estimate  $\hat{c}$  and  $k$ , we apply a minimum distance calibration procedure, which searches for a parameter constellation that minimizes the distance between the data and the model moments. The vector of distances  $\mathbf{g}(\hat{c}, k)$  between model and data moments

is made up of the following three elements:

$$g_1 = \chi(k, \hat{c}) - \chi_o = \hat{c}^k \left[ 1 - \frac{k}{k+1} \hat{c} \right] - \chi_o, \quad (22)$$

$$g_2 = \tilde{c}(k, \hat{c}) - \tilde{c}_o = \frac{k}{k+2} \hat{c}^{k+2} + \frac{k}{k+1} - \frac{k}{k+1} \hat{c}^{k+1} - \tilde{c}_o, \quad (23)$$

$$g_3 = v(k, \hat{c}) - v_o = \frac{k}{k+3} \hat{c}^{k+3} + \frac{k}{k+2} - \frac{k}{k+2} \hat{c}^{k+2} - [\tilde{c}(k, \hat{c})]^2 - v_o, \quad (24)$$

where  $\chi_o$  and  $\tilde{c}_o$  and  $v_o$  are the targeted moments observed in the data.<sup>8</sup>

The estimates of our structural parameters  $\hat{c}$  and  $k$  are as follows:

$$\{\hat{c}, k\} = \operatorname{argmin} \mathbf{g}(\hat{c}, k)' \mathbf{W} \mathbf{g}(\hat{c}, k), \quad (25)$$

*s.t.*

$$0 \leq \hat{c} \leq 1, \quad k > 1,$$

where  $\mathbf{W}$  is a  $3 \times 3$  weighting matrix. We adopt a diagonally weighted minimum distance approach, which requires that  $\mathbf{W}$  is diagonal and its diagonal elements are given by the variances of the data moments.<sup>9</sup>

To estimate the elasticity of substitution,  $\sigma$ , we can apply Eq. (6) for domestic producers and compute

$$\ln r^d(1-z) = \ln r^d(1) + (1-\sigma) \ln(1-z). \quad (26)$$

To recover  $\sigma$  and  $r_1$  we make use of four moment conditions. The first two are the moment conditions for the identification of the parameters in (26) in an OLS regression: the error term has to be uncorrelated with the covariates and its expected value must be equal to zero. We have

$$\begin{aligned} \zeta_1 &= E [\ln r^d - \ln r_1^d - (1-\sigma) \ln(1-z)] = 0, \\ \zeta_2 &= E [\{\ln r^d - \ln r_1^d - (1-\sigma) \ln(1-z)\} \ln(1-z)] = 0 \end{aligned} \quad (27)$$

Moreover, we exploit the time variation in our variables of interest, i.e. the length of the firm and firm revenue. Hence, we impose that the same moment conditions for identification hold for

---

<sup>8</sup>The model moments  $\tilde{c}$  and  $v$  are computed as  $\tilde{c} = \int_0^{\hat{c}} k(1-z)(1-z)^k + \int_{\hat{c}}^1 k(1-z)(1-z)^{k-1}$  and  $v = E(x^2) - [E(x)]^2 = \int_0^{\hat{c}} k(1-z)^2(1-z)^k + \int_{\hat{c}}^1 k(1-z)^2(1-z)^{k-1} - \tilde{c}^2$ .

<sup>9</sup>See Cameron and Trivedi (2005) for a detailed discussion of this approach.

the data in first differences. This enables us to exploit the following additional two equations:

$$\begin{aligned}\zeta_3 &= E [\Delta \ln r^d - (1 - \sigma)\Delta \ln(1 - z)] = 0, \\ \zeta_4 &= E [\{\Delta \ln r^d - (1 - \sigma)\Delta \ln(1 - z)\} \Delta \ln(1 - z)] = 0,\end{aligned}\tag{28}$$

where  $\Delta$  denotes the first difference of the data. Our estimates for  $\sigma$  and  $r_1$  thus solve:

$$\begin{aligned}\{\sigma, r_1\} &= \operatorname{argmin} \zeta(\sigma, r_1)' \mathbf{A} \zeta(\sigma, r_1), \\ & \text{s.t.} \\ & 1 \leq \sigma \leq k + 1, \quad r_1 > 0,\end{aligned}\tag{29}$$

where  $\mathbf{A}$  is a  $4 \times 4$  diagonal weighting matrix constructed as explained above. However, since we cannot observe the targets for  $\zeta_1$ - $\zeta_4$  in the data, but rather infer them from the theory underlying an OLS estimation, we have to construct matrix  $\mathbf{A}$  and set weights for the four moment conditions. We set the weight for the moment conditions in levels to 0.1 and choose a higher weight of 0.8 for the moment conditions in first differences. This guards us against biases in the estimation of  $\sigma$  due to unobserved time-invariant heterogeneity of firms.

In a next step, we aim to compute  $\kappa$ . For this purpose, we first determine total revenues of domestic firms:

$$R^d = M \int_0^1 r(c) f_c^d(c) dc = Mr(1) \left[ \frac{k}{k - \sigma + 2} \hat{c}^{k+2-\sigma} + \frac{k}{k - \sigma + 1} (1 - \hat{c}^{k+1-\sigma}) \right].\tag{30}$$

Accounting for Eq. (9), we can then compute

$$\frac{R^d}{R} = \frac{1 - \hat{c}^{k+1-\sigma} \left( 1 - \hat{c}^{\frac{k-\sigma+1}{k-\sigma+2}} \right)}{1 + \hat{c}^{k-\sigma+1} \left( 1 - \hat{c}^{\frac{k-\sigma+1}{k-\sigma+2}} \right) \left( \frac{\kappa^{1-\sigma}-1}{(1-\sigma)\ln \kappa} - 1 \right)}.\tag{31}$$

Considering the observed revenue ratio from the data and accounting for the parameter estimates of  $\sigma$ ,  $k$ , and  $\hat{c}$ , we can use this equation to recover a theory-consistent value of  $\kappa$ . Furthermore, we apply Eq. (7) to compute  $f$ , making use of  $\hat{c} = 1 - \hat{z}$  and consider the estimates for  $\sigma$ ,  $k$ ,  $\hat{c}$ , and  $r(1)$ . In a similar vein, we can make use of Eq. (15) to compute  $f_e$ . Finally, using the parameter estimates from above in  $\Gamma_2(\kappa, \hat{c}) = 0$ , we can recover a theory-consistent value of  $\tau L/L^*$ .

### 3.3 Estimation results

Table 4 gives an overview over the parameters estimated according the methods of moments (MOM) procedure outlined above. The upper panel reports the estimates for  $\hat{c}$  and  $k$ . The performance of our estimation can be evaluated by the difference between model and target moments reported in the third line. The middle panel reports the  $\sigma$  and  $r(1)$  estimates and the difference between the model and target moments. With these parameter estimates at hand, we can recover  $\kappa$ ,  $f$ ,  $f_e$ , and  $\tau L/L^*$  as reported in the lower panel of Table 4.

Solving problem (25) yields a local minimum at  $\hat{c} = 0.996, k = 1.653$ . The solution appears to be regular as measured by the gradient vector of the solver but the distance between the estimated model moments and the targeted moments is greater zero. The model precisely replicates the share of offshoring firms but it slightly deviates for the first moment of the distribution of  $1 - z$ . The weighting matrix gives highest weight to the observable share of offshoring firms and lower weights to the mean and variance of  $1 - z$ , which explains the larger distance between model and target for these moments. Solving problem (29) gives a local minimum at  $\sigma = 1.857$  and  $r(1) = 1,421,002$ . Again the solution appears to be regular as measured by the gradient vector of the solver and the difference between the estimated and targeted moments is in general fairly small. It is larger for the moments of the data in levels ( $\zeta_1$  and  $\zeta_2$ ) than the moments of the data in first differences ( $\zeta_3$  and  $\zeta_4$ ). This reflects that we choose higher weights on first difference targets than level targets in the minimization problem.

In a next step, we use the parameter estimates from the MOM regressions in Eqs. (7), (15), (30), and  $\Gamma_2(\kappa, \hat{c})$  to recover the labor cost differential  $\kappa$ , fixed costs  $f$ ,  $f_e$ , and the effective relative factor endowment  $\tau L/L^*$ . With a computed value of  $\kappa = 0.115$ , the cost saving from offshoring is sizable. Furthermore, the difference between the fixed costs of offshoring,  $f = 5,704.08$ , and the fixed costs of market entry,  $f_e = 3,265,730$ , is huge. This is because offshoring is attractive for German firms of all size and revenue categories, which requires low fixed costs of offshoring in our model. A computed value of  $\tau L/L^* = 0.522$  indicates that Germany has strong vertical linkages with other countries. However, treating Germany as a small economy seems not be warranted by this result.

At a first glance, a  $\sigma$  value of 1.857 might seem fairly low relative to the estimates reported in other empirical studies. However, the value is in line with the results reported by Broda and Weinstein (2006) who estimate the elasticity of substitution between consumer goods using

Table 4: Results for the methods of moments estimation

<i>Estimated parameters: <math>\hat{c}</math>, <math>k</math></i>						
<i>Targeted moments: Share of non-offshoring firms, mean and variance of <math>c = 1 - z</math></i>						
	$\hat{c}$	$k$	$\chi$	$\bar{c}$	$\text{var}(c)$	
Estimates	0.996	1.653	0.384	0.555	0.016	
Targets			0.377	0.452	0.150	
Difference			0.007	0.103	-0.134	
<i>Estimated parameters: <math>\sigma</math> and <math>r(1)</math></i>						
<i>Targeted moments: OLS moments conditions</i>						
	$\sigma$	$r(1)$	$\zeta_1$	$\zeta_2$	$\zeta_3$	$\zeta_4$
Estimate	1.857	1,421,002	-0.170	-0.208	-0.031	0.015
Targets			0	0	0	0
Difference			-0.170	-0.208	-0.031	0.015
<i>Recovered parameters: <math>\kappa</math>, <math>f</math>, <math>f_E</math>, and <math>\tau L/L^*</math></i>						
	$\kappa$	$f$	$f_e$	$\tau L/L^*$		
Parameters	0.115	5,704,.08	3,265,730	0.522		

Notes: Columns 4-6 in the first and second line of the first panel report the moments obtained from the model (computed using the minimum distance estimates) and the respective moments from the data. The difference between the estimated and the targeted moments are given in the third line. Columns 4-7 in the first and second line of the second panel report the estimated and targeted moments of the OLS model in Eqs. (27) and (28), with the difference given in the third line.

U.S. import data. They document a sizable industry variation in the elasticity of substitution between consumer goods, with the median level of  $\sigma$  declining in the degree of aggregation and over time. For the period 1990 to 2001 Broda and Weinstein (2006) report a median estimate for  $\sigma$  of 2.2 when relying on SITC-3 industry classifications. This is fairly close to the  $\sigma$ -estimate for our one-sector economy. To check the robustness of our result, we have estimated  $\sigma$  relying on OLS, random-, and fixed-effects regressions as possible alternatives to the MOM approach. The respective results are listed in Table 5. Our preferred MOM approach has the advantage that we are able to take both the level and the changes into consideration when estimating  $\sigma$ , which significantly lowers the problem of an omitted variable bias relative to the OLS estimates, reported in the second column of Table 5. The random effects (RE) estimates in the third

column are likely to suffer from the same kind of bias, since the assumption underlying the RE model is that unobserved firm characteristics are random and uncorrelated with the regressors. A simple Hausman test gives chi-square=206.55, which does not support the hypothesis that the firm-specific effects are uncorrelated with the regressors and suggests instead that firm-level fixed effects (FE) are more adequate for our problem. The estimates for the FE model are reported in the fourth column of Table 5. The estimated coefficient is not significant at the 10%-level, but the resulting  $\sigma$  is statistically different from zero. The  $\sigma$ -value from the MOM estimation lies between the estimates of the FE, RE and OLS models. A further advantage of the MOM approach to estimate  $\sigma$  compared to the three alternatives in Table 5 is that this approach allows us to impose the constraint that  $\sigma - 1 < k$ , which must hold to ensure a finite value of average revenues in our model. Aside from the MOM, this restriction dictated by the model is fulfilled in the FE regression only.

Table 5: Alternative approaches for estimating  $\sigma$

<i>Estimated Model:</i>			
$\ln r^d(1 - z) = \ln r^d(1) + (1 - \sigma) \ln(1 - z)$			
Estimator	OLS	RE	FE
$\ln c = \ln(1 - z)$	-3.197*** (0.068)	-2.364*** (0.110)	-0.252 (0.163)
$\sigma$	4.197***	3.364***	1.252***
$r(1)$	177,349.27	372,411.48	1,966,731.7
R-squared	0.565	0.565	0.987
Observations	2239	2239	2239

Notes: Robust standard errors in parentheses, \* significant at 10%, \*\* significant at 5%, \*\*\* significant at 1%.

### 3.4 Assessing the model fit

To determine the goodness of fit of our model, we determine the economy-wide overlap from the data, relying on percentile information. This gives a value of 0.536. Evaluating Eq. (19) at the parameter estimates from Table 4, we compute an economy-wide overlap of 0.187 from our theoretical model. Hence, our model is capable of explaining about 35 percent of the overlap in the data-set. To shed further light on this issue, we rank firms according to their revenues



and measure the overlap for deciles of the revenues distribution. To avoid that our results are driven by just a few outliers in the data-set, we compute the share of offshoring firms within certain decile intervals (0–1, 1–2, ..., 8–9) and determine the decile-specific offshoring index by applying the formula in Eq. (17).<sup>10</sup> We can then contrast the thus constructed offshoring indices with the respective indices computed from our model, when using the parameter estimates from Table 4. We report the ‘observed’ and ‘computed’ overlap indices together with the difference of these measures in Table 6.<sup>11</sup>

Table 6: Model fit: overlap

Decile	Overlap		Difference
	observed	computed	
1	0.407	0.002	0.405
2	0.49	0.012	0.478
3	0.704	0.037	0.667
4	0.907	0.103	0.804
5	0.868	0.276	0.592
6	0.774	0.744	0.031
7	0.442	0.495	-0.053
8	0.466	0.11	0.355
9	0.452	0.026	0.426
Average	0.612	0.201	0.412

While our model captures the hump shape of the overlap in the data-set fairly well, it significantly underestimates the overlap of offshoring and non-offshoring firms at the two tails of the revenue distribution, and in particular for firms with revenues below the mean. The correlation coefficient between the observed and computed overlap amounts to 0.28. Taking stock, we can conclude that our theoretical model clearly improves the understanding about the overlap of offshoring and non-offshoring firms. But the model is too parsimonious, and can therefore not tell the whole story behind the observed overlap in the data.

<sup>10</sup>Of course, the respective equation determines the overlap as a function of marginal costs and not revenues. However, since our model gives a one-to-one matching of a firm’s position in the marginal cost distribution and its position in the revenue distribution, the offshoring index does not depend on whether we rank firms according to their marginal costs or revenues. The difference is that we can observe revenues but not costs in the data.

<sup>11</sup>In the appendix, we consider revenue percentiles instead of revenue deciles for constructing the overlap indices and report the respective results in Table B.4.

## 4 Quantifying the effect of ignoring the overlap in the data

To see how important it is to account for the overlap in the data, we consider a variant of our model, in which all firms that make the fixed cost investment  $f$  will actually start offshoring, i.e. we set the probability of firms drawing  $s$  from the unit interval equal to one. Delegating formal details on how to solve the respective model to the appendix, we can newly estimate parameters  $\hat{c}$  and  $k$ , when replacing Eqs. (20) and (21) by

$$f_c^d(c) = \begin{cases} 0 & \text{if } c \leq \hat{c} \\ kc^{k-1} & \text{if } c > \hat{c} \end{cases} \quad (20')$$

and

$$\chi = \hat{c}^k, \quad (21')$$

respectively. This gives the new parameter estimates in Table 7.<sup>12</sup> Contrasting these parameter estimates with those from Table 4, we can conclude that ignoring the overlap of offshoring and non-offshoring firms in our model does not exert a sizable effect on the estimated value of shape parameter  $k$  of the Pareto distribution of  $z$ , but it leads to a significant downward bias in the  $\hat{c}$  estimate of almost 47 percent.

Replacing Eqs. (7),  $\Gamma_2 = 0$ , (15), and (31) by

$$\sigma f = (1 - \hat{z})^{1-\sigma} r^d(1) \left[ \frac{\kappa^{1-\sigma} - 1}{(1 - \sigma) \ln \kappa} - 1 \right], \quad (7')$$

$$\Gamma_2(\kappa, \hat{c}) = \kappa_n \left\{ \frac{\sigma + 1}{\sigma - 1} + \frac{2\sigma}{\sigma - 1} \frac{(1 - \sigma) \ln \kappa}{\kappa^{1-\sigma} - 1} \left( \frac{1}{\hat{c}^{k-\sigma+1}} - 1 \right) \right\} - \frac{\tau L}{L^*} = 0,$$

$$r(1) = \sigma f \left[ \frac{f_e}{f} - \frac{\sigma - 1}{k - \sigma + 1} \hat{c}^k \right] \frac{k - \sigma + 1}{k}, \quad (15')$$

---

<sup>12</sup>We do not report the estimates for  $\sigma$  and  $r(1)$  in Table 7, because the estimates do not differ from those in Table 4.

Table 7: A model variant without overlap

<i>Estimated parameters: <math>\hat{c}</math>, <math>k</math></i>					
<i>Targeted moments: Share of non-offshoring firms, mean and variance of <math>c = 1 - z</math></i>					
	$\hat{c}$	$k$	$\chi$	$\tilde{c}$	$\text{var}(c)$
Estimates	0.529	1.525	0.384	0.555	0.016
Targets			0.379	0.483	0.154
Difference			0.005	0.072	-0.138
<i>Recovered parameters: <math>\kappa</math>, <math>f</math>, <math>f_e</math>, and <math>\tau L/L^*</math></i>					
	$\kappa$	$f$	$f_e$	$\tau L/L^*$	
Parameters	0.247	1,229,820	2,345,320	1.118	

Notes: Columns 4-6 in the first and second line of the first panel report the moments obtained from the model (computed using the minimum distance estimates) and the respective moments from the data. The difference between the estimated and the targeted moments are given in the third line.

and

$$\frac{R^d}{R} = \frac{1 - \hat{c}^{k+1-\sigma}}{1 - \hat{c}^{k-\sigma+1} \left(1 - \frac{1-\kappa^{1-\sigma}}{(\sigma-1)\ln \kappa}\right)}, \quad (31')$$

respectively, we can compute theory-consistent values of  $\kappa$ ,  $f$ ,  $f_e$ , and  $\tau L/L^*$  for the model variant without overlap, as reported in Table 7. From a comparison with the respective results in Table 4, we can conclude that the estimates of  $\kappa$ ,  $f$ , and  $\tau L/L^*$  are significantly higher in the model variant without overlap than in the benchmark model with overlap, whereas there is a downward bias in the estimate of  $f_e$  when ignoring the overlap of offshoring and non-offshoring firms in the data.

With the parameter estimates at hand, we can now compute the welfare effects of offshoring for the two model variants. Using per-capita income as the preferred welfare measure and accounting for Eq. (16), we can compute the percentage increase of real wages in Germany that is explained by the observed exposure to offshoring in the model variant with overlap according to

$$\Delta W = 100 \left\{ \left(1 + \frac{\kappa L^*}{\tau L}\right)^{\frac{1}{\sigma-1}} \left[1 - \frac{\hat{c}^k}{1 - \hat{c}} \left(\frac{\sigma - 1}{k - \sigma + 1} - \hat{c} \frac{\sigma - 2}{k - \sigma + 2}\right) \frac{f}{f_e}\right]^{\frac{1}{1-\sigma}} - 1 \right\}. \quad (32)$$

Evaluating Eq. (32) at the parameter estimates from Table 4 shows that the welfare gain from offshoring is sizable and amounts to  $\Delta W = 192.29$  percent. This effect is likely to exaggerate the true gains from offshoring for two reasons. On the one hand, the exposure to offshoring is stronger for the manufacturing than the non-manufacturing industries, and hence economy-wide welfare gains from offshoring should be lower than those for manufacturing industries only. On the other hand, getting access to offshoring shifts the source country from a closed to an open economy, with all forms of trade. Since Germany is a very open economy, this means a huge shock, leading to a sizable welfare gain in our setting.

In the model variant without overlap, the gain from the observed exposure to offshoring is given by

$$\Delta W = 100 \left[ \left( 1 + \frac{\kappa L^*}{\tau L} \right)^{\frac{1}{\sigma-1}} \left( 1 - \hat{c}^k \frac{\sigma-1}{k-\sigma+1} \frac{f}{f_e} \right)^{\frac{1}{1-\sigma}} - 1 \right]. \quad (33)$$

Evaluated at the parameter estimates from Table 7, we compute a welfare gain from offshoring in the magnitude of  $\Delta W = 77.95$  for Germany. Contrasting welfare effects of offshoring for the two model variants, we can conclude that ignoring the overlap of offshoring and non-offshoring firms in the data leads to a severe downward bias in the estimated welfare effects of offshoring of almost 60 percent.<sup>13</sup>

## 5 Counterfactual analysis

With the parameter estimates at hand, we can conduct counterfactual experiments to shed light on how changes in offshoring costs affect the variables of interest. In the following experiment, we focus on changes in the fixed offshoring cost  $f$  and shed light on the decomposition of welfare effects along the intensive and the extensive margin of offshoring. We associate the intensive margin with welfare changes for a given share of offshoring firms,  $\chi$ . According to Eq. (21), the intensive margin thereby corresponds to welfare changes due to adjustments in  $\kappa$  for a given  $\hat{c}$ . The extensive margin is then associated with the welfare changes due to adjustments in  $\hat{c}$  for a given  $\kappa$ . Beyond that we also aim to shed light on how changes in offshoring costs affect the

---

<sup>13</sup>When relying on the parameter estimates based on value added instead of revenues as reported in Tables B.1 and B.3, we compute a welfare gain from offshoring that amounts to 192.86 and 76.95 for the model variant with and without overlap, respectively. This gives a downward bias in the welfare effect of offshoring when ignoring the overlap, which is very close to the one reported above.

overlap of offshoring and non-offshoring firms in our model. Figure 4 summarizes the results from this experiment.

The total welfare effects of offshoring for different levels of  $f$  are captured by the blue curve. From Figure 3 we know that higher offshoring fixed costs lower the cost cutoff  $\hat{c}$  and render the investment into exploring the offshoring opportunities less attractive for firms. As illustrated by the blue curve in Figure 4, this has negative consequences for the source country of offshoring. The red curve illustrates the welfare changes due to adjustments at the intensive margin, keeping the share of offshoring firms  $\chi$  at the respective value from our data-set. From our theoretical analysis, it follows that higher offshoring fixed costs lower labor demand in the host country, triggering a reduction in  $\kappa$  (see Figure 3). Whereas a lower  $\kappa$  reduces the costs of foreign production and therefore raises domestic welfare *ceteris paribus*, there is a counteracting effect, because lower host country wages reduce total consumer demand and thus the mass of firms active in the source country. Consumers therefore lose because they have access to less varieties, and this effect dominates the cost-saving effect of a decline in  $\kappa$ .<sup>14</sup>

From the gradients of the two loci representing the welfare effects of higher offshoring costs, we can infer that the intensive margin of welfare adjustments is particularly strong at low levels of  $f$ , i.e. high levels of economic integration. In the limiting case of  $f = 0$ , all firms find it attractive to invest into the exploration of their offshoring opportunities, and in this case any further welfare gain – for instance, due to a reduction in the variable costs of offshoring – can only materialize due to adjustments at the intensive margin. In contrast, if  $f$  falls from infinity to a high finite value, there are no incumbent offshoring firms who can benefit from the cost reduction of offshoring, and hence it must be the extensive margin (in addition to a small expansion of consumer demand) that generates positive welfare effects in this case.

The black curve in Figure 4 captures the overlap of offshoring and non-offshoring firms for different values of offshoring fixed cost  $f$ . The curve confirms our theoretical insight that the overlap is small if the costs of offshoring is either high or low. If offshoring costs are high only few very productive firms make the investment into exploring their offshoring opportunities and

---

<sup>14</sup>If  $\hat{c}$  stays constant, the fall in  $M$  following an increase in  $f$  is more pronounced than in an otherwise symmetric scenario, in which  $\hat{c}$  responds to the increase in  $f$ . The reason is that the offshoring decision is distorted if  $\hat{c}$  is held constant, with an excessive share of firms investing into the exploration of their offshoring opportunities. The additional fixed costs from excessive offshoring renders firms entry less attractive and therefore reduces the mass of active firms relative to a scenario, in which firms adjust their offshoring decision after an increase in the fixed cost of offshoring.

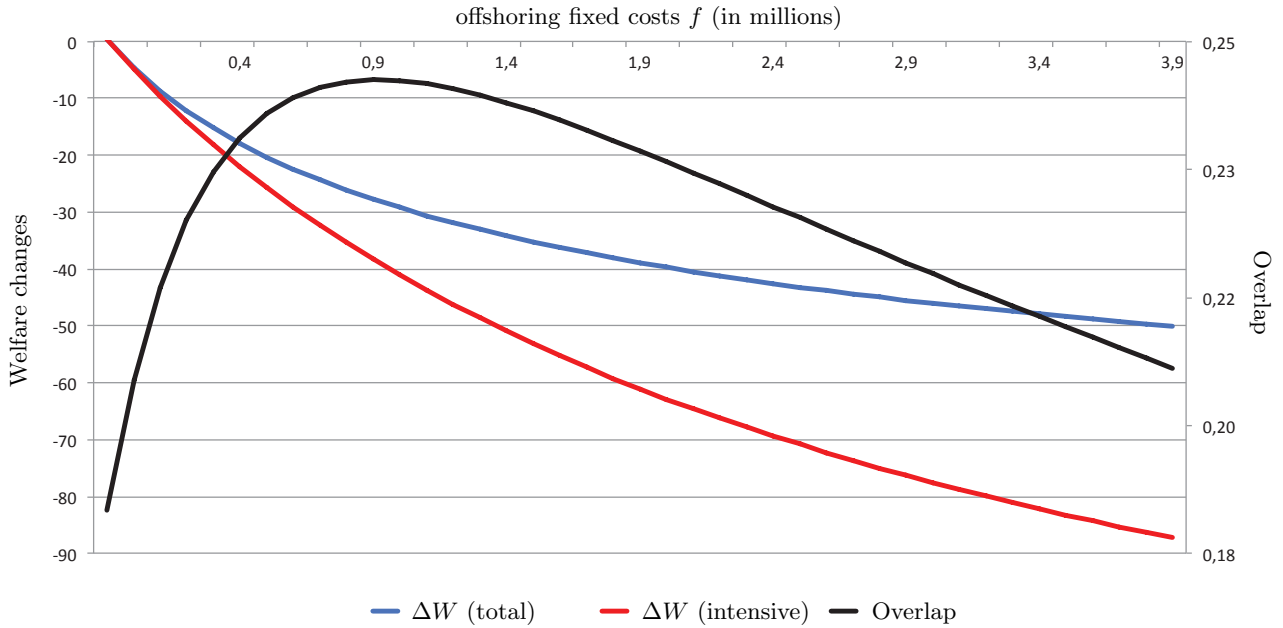


Figure 4: Offshoring costs, welfare and overlap

among these firms the probability of offshoring is large, which implies a small overlap. For low offshoring costs, the marginal firm that is indifferent between making and not making the investment of  $f$  is one with a low probability of offshoring, and for firms of this cost category the overlap is therefore small. A further decline in  $f$  must thus reduce the economy-wide overlap. For intermediate cost levels, the overlap reaches a maximum of about 24.5 percent, which is still lower than the overlap observed in our data-set.

## 6 Concluding remarks

This paper sets up a model of heterogeneous firms, in which firms differ in the number of tasks they have to 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 depend negatively on the number of tasks performed and the share of tasks offshored. Offshoring is subject to fixed and variables trade costs, and not all firms find it attractive to make the investment into offshoring. This gives a model of

heterogeneous 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 and the number of tasks performed in the production process. We introduce an index to measure the overlap of offshoring and non-offshoring firms and show that this index is non-monotonic in the costs of offshoring.

In an empirical exercise, we estimate key parameters of the model, using firm-level data from Germany. We use a method of moments approach, in which parameters are estimated by minimizing the distance between moments of the model and the data. Based on these parameter estimates, we show that our model can explain about 35 percent of the overlap in the data. Furthermore, we quantify the bias in welfare estimates from ignoring the overlap of offshoring and non-offshoring firms as observed in our data-set. Thereby, the respective bias turns out to be huge. Ignoring the overlap in the data leads to a severe downward bias in the welfare gains of offshoring predicted by our model of almost 60 percent. In a counterfactual analysis, we shed light on the welfare effects of changes in offshoring costs and offer a decomposition into adjustments along the extensive and intensive margin of offshoring. We also show that the overlap is hump-shaped in the costs of offshoring.

Whereas we think that this paper provides a useful starting point for studying overlap in models of heterogeneous firms, it is clear that the parsimonious structure lowers the ability of our model to capture the true empirical pattern of overlap. For instance, the model in its present form underestimates the overlap of offshoring and non-offshoring firms at the lower and upper tail of the revenue distribution. The main reason for this is the underlying assumption regarding the link of the number of tasks performed and the share of offshorable tasks in the production of goods. Choosing a more flexible structure for the distributions of the two random variables, while lowering analytical tractability, would probably help improving the explanatory power of the model regarding the overlap in the data. Going in this direction would therefore be a worthwhile task 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.

- 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.
- 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.
- 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. AND J. B. JENSEN (1995): “Exporters, Jobs and Wages in US Manufacturing: 1976-1987,” Brooking Papers on Economics Activity, Microeconomics, 67–119.
- 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.
- BRODA, C. AND D. E. WEINSTEIN (2006): “Globalization and the Gains from Variety,” Quarterly Journal of Economics, 121, 541–585.
- CAMERON, A. C. AND P. K. TRIVEDI (2005): Microeconometrics – Methods and Applications, Cambridge University Press.
- CHANG, H. AND C. VAN MARREWIJK (2011): “Firm Heterogeneity and Development: Evidence from Latin American Countries,” Tjalling C. Koopmans Research Institute Discussion Paper 11-14.
- DAVIS, D. R. AND J. HARRIGAN (2011): “Good Jobs, Bad Jobs, and Trade Liberalization,” Journal of International Economics, 84, 26–36.
- DHINGRA, S. AND J. MORROW (2013): “Monopolistic Competition and Optimum Product Diversity under Firm Heterogeneity,” Unpublished Manuscript, London School of Economics.
- EGGER, H., U. KREICKEMEIER, AND J. WRONA (2013): “Offshoring Domestic Jobs,” CESifo Working Paper 4083, CESifo.
- 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 data  $\frac{1}{4}$ , 01/2008.
- GREENAWAY, D., J. GULLSTRAND, AND R. KNELLER (2005): “Exporting May Not Always Boost Firm Productivity,” Review of World Economics, 141, 561–582.
- HARRIGAN, J. AND A. RESHEF (2015): “Skill Biased Heterogeneous Firms, Trade Liberalization, and the Skill Premium,” Canadian Journal of Economics, 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.



- HUMMELS, D. AND P. J. KLENOW (2005): “The Variety and Quality of a Nation’s Exports.” American Economic Review, 95, 704 – 723.
- 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 (2014): “On the Heterogeneous Employment Effects of Offshoring – Identifying Productivity and Downsizing Effects,” Economic Inquiry, forthcoming.
- NELSON, R. B. (2007): An Introduction to Copulas, Springer Series in Statistics, Springer, 2nd ed.
- 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.
- SCHRÖDER, P. J. H. AND A. SØRENSEN (2012): “Second Thoughts on the Exporter Productivity Premium,” Canadian Journal of Economics, 45, 1310–1331.
- SPITZ-OENER, A. (2006): “Technical Change, Job Tasks and Rising Educational Demands: Looking Outside the Wage Structure,” Journal of Labor Economics, 24, 235–270.

## A Theoretical appendix

### A.1 Derivation of Eq. (8)

Let us define  $b(z) = 1 - z$ . Then, the probability of  $b \leq \bar{b}$ ,  $P_b(b \leq \bar{b})$  equals the probability of  $z \geq 1 - \bar{b}$ ,  $P_z(z \geq 1 - \bar{b})$ . Accounting for

$$P_z(z \geq 1 - \bar{b}) = 1 - P_z(z \leq 1 - \bar{b}) = 1 - \int_0^{1-\bar{b}} f_z(z) dz, \quad (\text{A.1})$$

we can compute  $P_b(b \leq \bar{b}) = 1 - \bar{b}^k$ . The cumulative distribution function of  $b$  is therefore given by  $F_b(b) = \bar{b}^k$ . Since  $c = b$  if  $z \leq \hat{z}$  and thus  $c \geq \hat{c}$ , the third segment of the productivity density function of  $c$  is given by  $f_c(c) = kc^{k-1}$ .

To determine the probability density function of  $c$  for interval  $c \leq \hat{c}$ , we can first note that the probability for  $1 - \hat{b} \leq z \leq 1 - \bar{b}$  is given by  $P_z(1 - \hat{b} \leq z \leq 1 - \bar{b}) = k \int_{1-\hat{b}}^{1-\bar{b}} (1-z)^{k-1} dz$  and disentangling non-offshoring firms (superscript  $d$ ) from offshoring firms (superscript  $o$ ), we can write

$$P_z(1 - \hat{b} \leq z \leq 1 - \bar{b}) = \underbrace{k \int_{1-\hat{b}}^{1-\bar{b}} (1-z)^k dz}_{P_z^d(\cdot)} + \underbrace{k \int_{1-\hat{b}}^{1-\bar{b}} z(1-z)^{k-1} dz}_{P_z^o(\cdot)}. \quad (\text{A.2})$$

Solving the integrals establishes

$$P_z^d(1 - \hat{b} \leq z \leq 1 - \bar{b}) = -\frac{k}{k+1} \left[ \bar{b}^{k+1} - \hat{b}^{k+1} \right] \quad (\text{A.3})$$

$$P_z^o(1 - \hat{b} \leq z \leq 1 - \bar{b}) = -\frac{1}{1+k} \left[ \bar{b}^k (1 + k(1 - \bar{b})) - \hat{b}^k (1 + k(1 - \hat{b})) \right] \quad (\text{A.4})$$

Summing up and adding  $P_z(z \leq 1 - \hat{b}) = 1 - \hat{b}^k$ , we can compute the probability of  $z \leq (1 - \bar{b})$ :

$$P_z(z \leq 1 - \bar{b}) = 1 - \frac{k}{k+1} \bar{b}^{k+1} - \frac{1}{k+1} \bar{b}^k [1 + k(1 - \bar{b})] = \frac{1}{k+1} (1 - \bar{b}^k). \quad (\text{A.5})$$

In view of  $P_b(b \leq \bar{b}) = 1 - P_z(z \leq 1 - \bar{b})$ , we can express the cumulative distribution function of  $b$  as  $F_b(b) = F_b^d(b) + F_b^o(b)$ , with  $F_b^d(b) = [k/(k+1)]b^{k+1}$  and  $F_b^o = b^k - [k/(k+1)]b^{k+1}$  in the relevant interval. This establishes the probability density function  $f_b(b) = f_b^d(b) + f_b^o(b)$ , with  $f_b^d(b) \equiv kb^k$  and  $f_b^o(b) \equiv kb^{k-1}(1-b)$ , respectively. For purely domestic producers, we have  $c = b$ , and can thus write  $f_c^d(c) = kc^k$ . For offshoring firms, things are different, because  $\kappa < 1$  establishes  $c = b\kappa^s < b$ . Let us define  $a \equiv \kappa^s$ . We can compute  $s = \ln a / \ln \kappa$ , and hence can write  $P_a(a \leq \bar{a}) = P_s(s \geq s(\bar{a})) = 1 - P_s(s \leq s(\bar{a}))$ . Accounting for  $P_a(a \leq \bar{a}) = 1 - \int_0^{s(\bar{a})} 1 ds$  allows us to compute  $P_a(a \leq \bar{a}) = 1 - \ln \bar{a} / \ln \kappa$ . The probability density function of  $a$  is then given by  $f_a(a) = -1/(a \ln \kappa)$ .

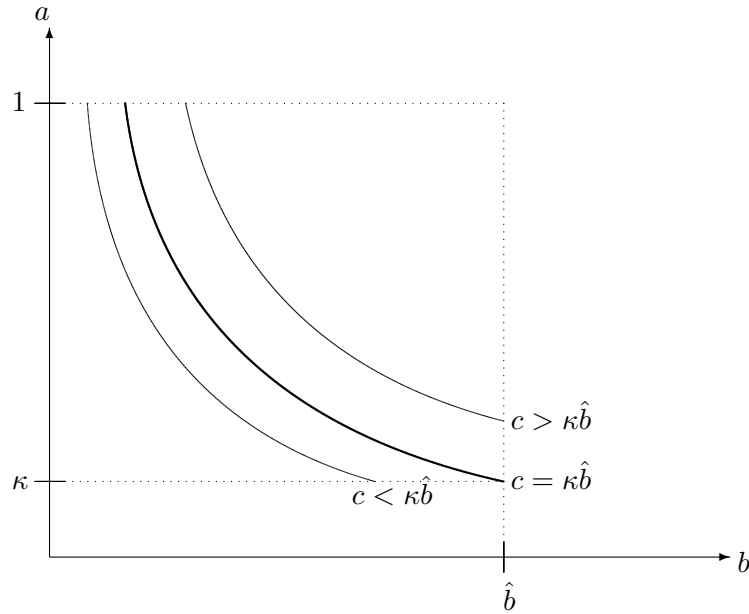


Figure A.1: Iso- $c$  lines in  $(b, a)$ -space

We now have all necessary ingredients and can compute the probability density function of

$c$  for those firms drawing  $s$  from the unit interval according to

$$f_c^o(c) = \int_{b \in B} f_b^2(b) f_a\left(\frac{c}{b}\right) \left| \frac{1}{b} \right| db = -\frac{1}{c \ln \kappa} \int_{b \in B} k b^{k-1} (1-b) db, \quad (\text{A.6})$$

where  $B$  is the set of feasible  $b$ 's. To determine the explicit bounds of the integral, we can look at Figure A.1. There, we see 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

$$\begin{aligned} f_c^{o1}(c) &= -\frac{k}{c \ln \kappa} \int_c^{c/\kappa} (b^{k-1} - b^k) db \\ &= -\frac{1}{\ln \kappa} \left\{ c^{k-1} \left[ \left( \frac{1}{\kappa} \right)^k - 1 \right] - \frac{k c^k}{k+1} \left[ \left( \frac{1}{\kappa} \right)^{k+1} - 1 \right] \right\}. \end{aligned} \quad (\text{A.7})$$

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

$$\begin{aligned} f_c^{o2}(c) &= -\frac{k}{c \ln \kappa} \int_c^{\hat{b}} (b^{k-1} - b^k) db \\ &= -\frac{1}{\ln \kappa} \left\{ c^{k-1} \left[ \left( \frac{\hat{b}}{c} \right)^k - 1 \right] - \frac{k c^k}{k+1} \left[ \left( \frac{\hat{b}}{c} \right)^{k+1} - 1 \right] \right\}. \end{aligned} \quad (\text{A.8})$$

Replacing  $\hat{b}$  by  $\hat{c}$ , and summing up  $f_c^d(c)$  and  $f_c^o(c)$  for the two parameter domains gives the first and the second segment of the probability density function in Eq. (8). This completes the proof. *QED*

## A.2 Derivation of Eq. (9)

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

$$\begin{aligned} R_1 &= Mr(1) \int_0^{\kappa \hat{c}} c^{1-\sigma} f_c(c) dc \\ &= Mr(1) \int_0^{\kappa \hat{c}} c^{1-\sigma} \left\{ k c^k - \frac{1}{\ln \kappa} \left\{ c^{k-1} \left[ \left( \frac{1}{\kappa} \right)^k - 1 \right] - \frac{k c^k}{k+1} \left[ \left( \frac{1}{\kappa} \right)^{k+1} - 1 \right] \right\} \right\} dc. \end{aligned} \quad (\text{A.9})$$

Solving for the integral, gives

$$R_1 = Mr(1) \left\{ \frac{k}{k-\sigma+2} (\kappa \hat{c})^{k-\sigma+2} - \frac{\kappa^{-k} - 1}{\ln[\kappa]} \frac{1}{k-\sigma+1} (\kappa \hat{c})^{k-\sigma+1} \right\}$$

$$+ \frac{\kappa^{-(k+1)} - 1}{\ln[\kappa]} \frac{k}{k+1} \frac{1}{k-\sigma+2} (\kappa\hat{c})^{k-\sigma+2} \Big\}. \quad (\text{A.10})$$

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

$$\begin{aligned} R_2 &= Mr(1) \int_{\kappa\hat{c}}^{\hat{c}} c^{1-\sigma} f_c(c) dc \\ &= Mr(1) \int_{\kappa\hat{c}}^{\hat{c}} c^{1-\sigma} \left\{ kc^k - \frac{1}{\ln \kappa} \left\{ c^{k-1} \left[ \left( \frac{\hat{c}}{c} \right)^k - 1 \right] - \frac{kc^k}{k+1} \left[ \left( \frac{\hat{c}}{c} \right)^{k+1} - 1 \right] \right\} \right\} dc. \end{aligned} \quad (\text{A.11})$$

Solving for the integral establishes

$$\begin{aligned} R_2 &= Mr(1) \left\{ \left( 1 - \kappa^{k-\sigma+2} \right) \frac{k\hat{c}^{k-\sigma+2}}{k-\sigma+2} - \frac{1 - \kappa^{1-\sigma}}{\ln[\kappa]} \frac{\hat{c}^{k-\sigma+1}}{\sigma-1} + \frac{1 - \kappa^{k-\sigma+1}}{\ln[\kappa]} \frac{\hat{c}^{k-\sigma+1}}{k-\sigma+1} \right. \\ &\quad \left. + \frac{\kappa^{1-\sigma} - 1}{\ln[\kappa]} \frac{k\hat{c}^{k-\sigma+2}}{(\sigma-1)(k+1)} - \frac{1 - \kappa^{k-\sigma+2}}{\ln[\kappa]} \frac{k\hat{c}^{k-\sigma+2}}{(k+1)(k-\sigma+2)} \right\}. \end{aligned} \quad (\text{A.12})$$

Finally, for the first segment, we obtain

$$\begin{aligned} R_1 &= Mr(1) \int_{\hat{c}}^1 c^{1-\sigma} f_c(c) dc = Mr(1) \int_{\hat{c}}^1 c^{1-\sigma} kc^{k-1} dc \\ &= Mr(1) \left[ \frac{k}{k-\sigma+1} - \frac{k}{k-\sigma+1} \hat{c}^{k-\sigma+1} \right]. \end{aligned} \quad (\text{A.13})$$

Total revenues in Eq. (9) 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

$$\alpha(\kappa) = \frac{\kappa^{1-\sigma} - 1}{(1-\sigma)\ln \kappa} - 1, \quad (\text{A.14})$$

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

$$\alpha'(\kappa) = \frac{\hat{\alpha}(\kappa)}{(1-\sigma)[\ln \kappa]^2 \kappa^\sigma}, \quad \hat{\alpha}(\kappa) \equiv (1-\sigma)\ln \kappa + \kappa^{\sigma-1} - 1. \quad (\text{A.15})$$

Accounting for  $\lim_{\kappa \rightarrow 0} \hat{\alpha}(\kappa) = \infty$ ,  $\hat{\alpha}(1) = 0$ , 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

$$\frac{\partial \Gamma_1(\cdot)}{\partial \kappa} = \left\{ \frac{\hat{b}^k}{1 - \hat{b}} \left[ \frac{\sigma - 1}{k - \sigma + 1} - \hat{b} \frac{\sigma - 2}{k - \sigma + 2} \right] - \frac{f_e}{f} \right\} \alpha'(\kappa) \quad (\text{A.16})$$

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  $\hat{b}$  yields

$$\begin{aligned} \frac{\partial \Gamma_1(\cdot)}{\partial \hat{c}} &= \frac{(\sigma - 1)\hat{c}^{\sigma-2}(1 - \hat{c}) + \hat{c}^{\sigma-1}}{(1 - \hat{c})^2} \frac{k}{k - \sigma + 1} + \frac{\hat{c}^k}{(1 - \hat{c})^2} \left[ \frac{\sigma - 1}{k - \sigma + 1} - \hat{c} \frac{\sigma - 2}{k - \sigma + 2} \right] \alpha(\kappa) \\ &\quad + \frac{\hat{c}^{k-1}}{1 - \hat{c}} \left[ \frac{k(\sigma - 1)}{k - \sigma + 1} - \hat{c} \frac{(k + 1)(\sigma - 2)}{k - \sigma + 2} \right] \alpha(\kappa) \end{aligned}$$

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  $k^2 > (\sigma - 2)(k - \sigma + 1)$  that  $k(\sigma - 1)/(k - \sigma + 1) > (k + 1)(\sigma - 2)/(k - \sigma + 2)$  and this is sufficient for the third term to be positive. This implies  $\partial \Gamma_1(\cdot)/\partial \hat{b} > 0$ . Putting together, we have shown

$$\left. \frac{d\hat{c}}{d\kappa} \right|_{\Gamma_1(\cdot)=0} = - \frac{\partial \Gamma_1/\partial \kappa}{\partial \Gamma_1/\partial \hat{c}} < 0. \quad (\text{A.17})$$

As noted in the main text,  $\Gamma_1(\cdot) = 0$  gives  $\hat{c}$  as an implicit function of  $\kappa$ . If  $\kappa$  goes to one,  $\alpha(\kappa)$  becomes zero, and hence  $\hat{c}$  must fall to zero in order to restore  $\Gamma_1(\cdot) = 0$ . In contrast, if  $\kappa$  falls to zero,  $\alpha$  goes to infinity, and hence  $\hat{c}$  must increase to  $\hat{c}_1$  in order to restore  $\Gamma_1(\cdot) = 0$ . This completes the formal discussion on the properties of *OC*. *QED*

#### A.4 Properties of the labor market constraint

To see whether LC establishes a positive or negative link between  $\kappa$  and  $\hat{c}$ , we can first look at the properties of  $\beta(\kappa) \equiv (1 - \sigma) \ln \kappa / (\kappa^{1-\sigma} - 1)$ . Differentiating  $\beta(\kappa)$  gives

$$\beta'(\kappa) = - \frac{(\sigma - 1)\hat{\beta}(\kappa)}{\kappa(\kappa^{\sigma-1} - 1)^2}, \quad \hat{\beta}(\kappa) \equiv \kappa^{1-\sigma} [1 - (1 - \sigma) \ln \kappa] - 1. \quad (\text{A.18})$$

Noting that  $\lim_{\kappa \rightarrow 0} \hat{\beta}(\kappa) = -\infty$ ,  $\hat{\beta}(1) = 0$ , and  $\hat{\beta}'(\kappa) = -(\sigma - 1)^2 \kappa^{-\sigma} \ln \kappa > 0$ , it is immediate that  $\beta'(\kappa) > 0$  holds for all possible  $\kappa$ . Thereby,  $\beta(\kappa)$  increases from a minimum value 0 if  $\kappa = 0$  to a maximum value of one at  $\kappa = 1$ . The positive sign of  $\beta'(\kappa) > 0$  establishes  $\partial \Gamma_2(\cdot)/\partial \kappa > 0$ .

In a second step, we can look at the properties of  $\gamma(\hat{c}) = \hat{c}^{k-\sigma+1} [1 + (1 - \hat{c})(k - \sigma + 1)]$ . Differentiating  $\gamma(\hat{c})$  gives  $\gamma'(\hat{c}) = (k - \sigma + 2)(k - \sigma + 1)\hat{c}^{k-\sigma}(1 - \hat{c}) > 0$ . Hence,  $\gamma(\hat{c})$  increases from a minimum value of zero at  $\hat{c} = 0$  to a maximum value of 1 at  $\hat{c} = 1$ . The positive sign of  $\gamma'(\hat{c})$  establishes  $\partial \Gamma_2(\cdot)/\partial \hat{c} < 0$ . Applying the implicit function theorem to  $\Gamma_2(\cdot) = 0$  therefore

establishes a positive link between  $\kappa$  and  $\hat{c}$ :

$$\left. \frac{d\kappa}{d\hat{c}} \right|_{\Gamma_2(\cdot)=0} = -\frac{\partial\Gamma_2/\partial\hat{c}}{\partial\Gamma_2/\partial\kappa} > 0. \quad (\text{A.19})$$

$\Gamma_2(\cdot) = 0$  determines  $\kappa$  as implicit function of  $\hat{c}$ . If  $\hat{c} = 0$ ,  $\gamma(\hat{c})$  falls to zero, and hence  $\kappa$  must fall to zero as well in order to establish  $\beta = 0$  and to restore  $\Gamma_2(\cdot) = 0$ . Furthermore, if  $\mu < 0$  (with  $\mu$  defined in the main text),  $\kappa$  is smaller than one for any  $\hat{c} \in [0, 1]$ . If  $\hat{c}$  goes to one,  $\gamma$  also reaches a value of 1, and in this case  $\kappa$  must increase to  $\kappa_2$  in order to restore  $\Gamma_2(\cdot) = 0$ . Things are different if  $\mu > 0$ . In this case,  $\kappa$  reaches a maximum value of one if  $\hat{c} = \hat{c}_2$ . This completes the formal discussion on the properties of  $LC$ . *QED*

### A.5 The impact of $\tau$ and $f$ on $W$

For a given global income,  $L + w^*L^*$ ,  $W = P^{-1}$  increases in  $\hat{c}$ , and hence it declines in  $\tau$  and  $f$ . Furthermore, from  $\Gamma_2(\cdot) = 0$ , we can infer

$$w^*L^* = L \left\{ \frac{\sigma + 1}{\sigma - 1} + \frac{2\sigma}{\sigma - 1} \beta(\kappa) \left[ \frac{k - \sigma + 2}{\gamma(\hat{c})} - 1 \right] \right\}^{-1} \quad (\text{A.20})$$

Accounting for  $\beta'(\kappa) > 0$  and  $\gamma'(\hat{c}) > 0$  from above and considering  $d\kappa/d\tau > 0$ ,  $d\hat{c}/d\tau < 0$  from Figure 3, we can safely conclude that  $d(w^*L^*)/d\tau < 0$ . Furthermore, considering  $w^*L^* = \kappa L^*/\tau$  and accounting for  $d\kappa/df < 0$  from Figure 3, it follows that  $d(w^*L^*)/df < 0$ . Putting together, we can safely conclude that  $W = P^{-1}$  decreases in  $f$  and  $\tau$ . This completes the proof. *QED*

### A.6 Proof of a hump shape of $O(c)$

We can define

$$a(c) \equiv \begin{cases} \frac{1}{c} \left[ \left( \frac{1}{\kappa} \right)^k - 1 \right] - \frac{k}{k+1} \left[ \left( \frac{1}{\kappa} \right)^{k+1} - 1 \right] & \text{if } c \leq \kappa\hat{c} \\ \frac{1}{c} \left[ \left( \frac{\hat{c}}{c} \right)^k - 1 \right] - \frac{k}{k+1} \left[ \left( \frac{\hat{c}}{c} \right)^{k+1} - 1 \right] & \text{if } c \in (\kappa\hat{c}, \hat{c}] \end{cases} \quad (\text{A.21})$$

with  $a'(c) < 0$ . Since in view of Eqs. (17) and (18), we can write

$$1 - 2\frac{kc^k}{f_c(c)} = \frac{a(c) + k \ln \kappa}{a(c) - k \ln \kappa} \quad (\text{A.22})$$

it follows that  $1 - 2kc^k/f_c(c)$  decreases in  $c$  from a maximum level of one at  $c = 0$  to a minimum level of  $-1$  at  $c = \hat{c}$ . This proves the hump-shaped pattern of  $O(c)$ . *QED*

## A.7 A model variant without overlap

To remove the overlap from the model we just have to set the probability of offshoring when making the  $f$  investment equal to one. Whereas this modification does not affect Eqs. (1)-(6), it changes indifference condition (7), which now reads:

$$\sigma f = (1 - \hat{z})^{1-\sigma} r^d(1) \left[ \frac{\kappa^{1-\sigma} - 1}{(1 - \sigma) \ln \kappa} - 1 \right]. \quad (\text{A.23})$$

To determine the probability density function of  $c$ , we can follow the steps in Appendix A.1 and obtain

$$f_c(c) = \begin{cases} -\frac{1}{\ln \kappa} c^{k-1} \left[ \left(\frac{1}{\kappa}\right)^k - 1 \right] & \text{if } c \leq \kappa \hat{c} \\ -\frac{1}{\ln \kappa} c^{k-1} \left[ \left(\frac{\hat{c}}{c}\right)^k - 1 \right] & \text{if } c \in (\kappa \hat{c}, \hat{c}] \\ kc^{k-1} & \text{if } c > \hat{c} \end{cases} \quad (\text{A.24})$$

instead of Eq. (8). In a next step, we can determine aggregate revenues:

$$R = M \int_0^1 r(c) f_c(c) dc = Mr(1) \frac{k}{k - \sigma + 1} \left[ 1 - \hat{c}^{k-\sigma+1} \left( 1 - \frac{1 - \kappa^{1-\sigma}}{\ln \kappa} \frac{1}{\sigma - 1} \right) \right] \quad (\text{A.25})$$

Using this in the free entry condition, we obtain the modified *offshoring indifference condition*:

$$\Gamma_1(\hat{c}, \kappa) \equiv \frac{k}{k - \sigma + 1} \hat{c}^{\sigma-1} \left[ \frac{(1 - \sigma) \ln \kappa}{\kappa^{1-\sigma} - 1 - (1 - \sigma) \ln \kappa} + \frac{\sigma - 1}{k} \hat{c}^{k-\sigma+1} \right] - \frac{f_e}{f} = 0, \quad (\text{A.26})$$

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

To get a second link between  $\hat{c}$  and  $\kappa$  we can make use of Eq. (11) and compute a modified *labor market condition*. Following the derivation steps from the main text, this gives:

$$\Gamma_2(\kappa, \hat{c}) \equiv \kappa \left\{ \frac{\sigma + 1}{\sigma - 1} + \frac{2\sigma}{\sigma - 1} \frac{(1 - \sigma) \ln \kappa}{\kappa^{1-\sigma} - 1} \left( \frac{1}{\hat{c}^{k-\sigma+1}} - 1 \right) \right\} - \frac{\tau L}{L^*} = 0, \quad (\text{A.27})$$

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

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

$$r(1) = \sigma f \left[ \frac{f_e}{f} - \frac{\sigma - 1}{k - \sigma + 1} \hat{c}^k \right] \frac{k - \sigma + 1}{k} \quad (\text{A.28})$$

and

$$P^{-1} = \frac{\sigma - 1}{\sigma} \left\{ \frac{L + w^* L^*}{\sigma f} \left[ \frac{f_e}{f} - \frac{\sigma - 1}{k - \sigma + 1} \hat{c}^k \right]^{-1} \frac{k}{k - \sigma + 1} \right\}^{\frac{1}{\sigma-1}}. \quad (\text{A.29})$$

With these insights at hand, we can finally solve for Eqs. (20'), (21'), and (31'). This completes the proof. *QED*



# B Graphical and data appendix

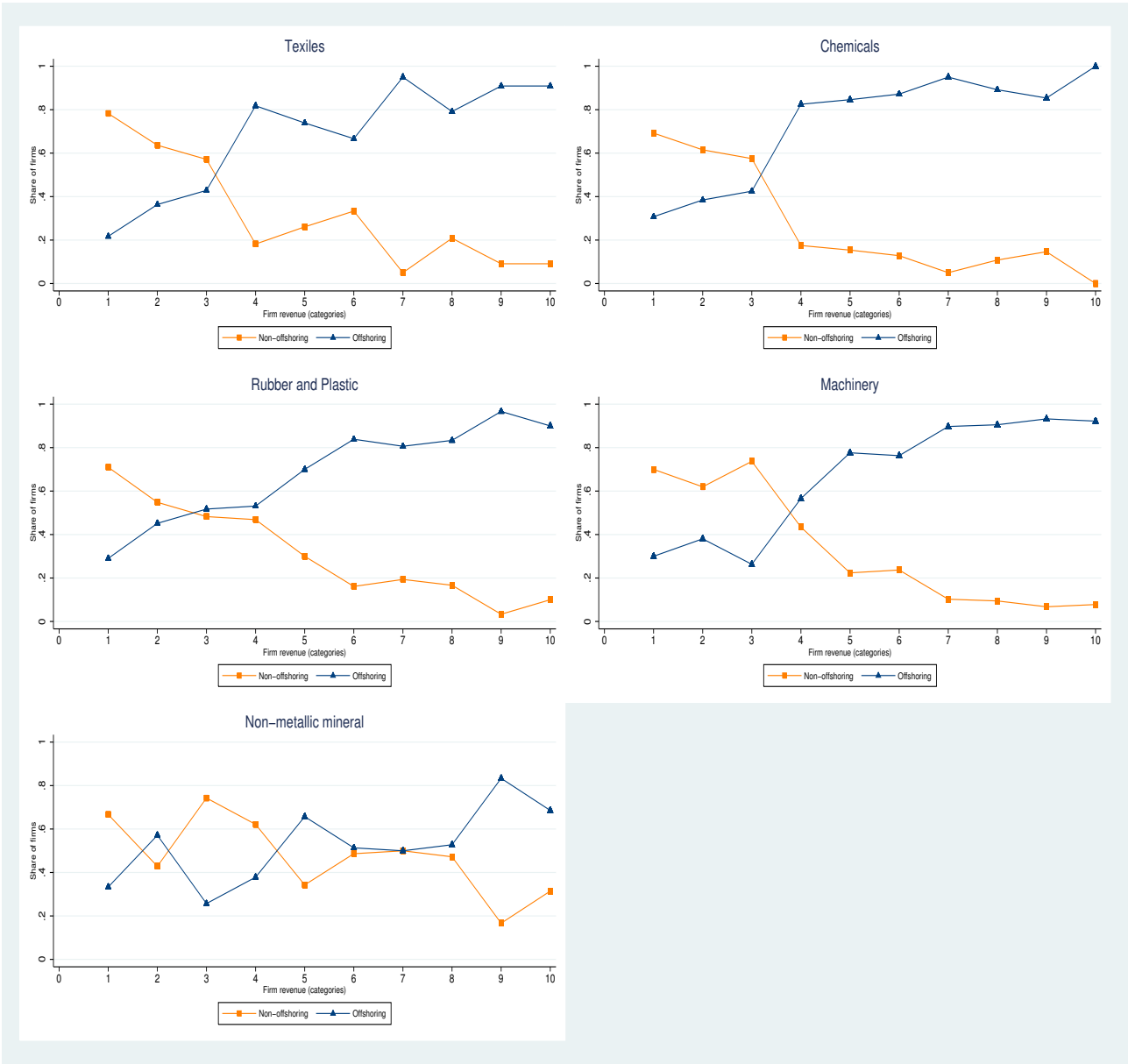


Figure B.2: Share of offshoring firms in different revenue categories at sector level

Table B.1: Results for the methods of moments estimation – value added

<i>Estimated parameters: <math>\hat{c}</math>, <math>k</math></i>						
<i>Targeted moments: Share of non-offshoring firms, mean and variance of <math>c = 1 - z</math></i>						
	$\hat{c}$	$k$	$\chi$	$\tilde{c}$	$\text{var}(c)$	
Estimates	0.996	1.653	0.384	0.555	0.016	
Targets			0.377	0.452	0.150	
Difference			0.007	0.103	-0.134	
<i>Estimated parameters: <math>\sigma</math> and <math>r(1)</math></i>						
<i>Targeted moments: OLS moments conditions</i>						
	$\sigma$	$r(1)$	$\zeta_1$	$\zeta_2$	$\zeta_3$	$\zeta_4$
Estimate	1,826	614,440.736	-0.156	-0.191	-0.020	0.013
Targets			0	0	0	0
Difference			-0.156	-0.191	-0.020	0.013
<i>Recovered parameters: <math>\kappa</math>, <math>f</math>, <math>f_E</math>, and <math>\tau L/L^*</math></i>						
	$\kappa$	$f$	$f_e$	$\tau L/L^*$		
Parameters	0.111	2,431.3	1,348,560	0.521		

Notes: Columns 4-6 in the first and second line of the first panel report the moments obtained from the model (computed using the minimum distance estimates) and the respective moments from the data. The difference between the estimated and the targeted moments are given in the third line. Columns 4-7 in the first and second line of the second panel report the estimated and targeted moments of the OLS model in Eqs. (27) and (28), with the difference given in the third line.

Table B.2: Alternative approaches for estimating  $\sigma$  – value added

<i>Estimated Model:</i>			
$\ln r^d(1 - z) = \ln r^d(1) + (1 - \sigma) \ln(1 - z)$			
Estimator	OLS	RE	FE
$\ln c = \ln(1 - z)$	-3.022*** (0.077)	-2.687*** (0.096)	-0.319 (0.340)
$\sigma$	4.022***	3.687***	1.318***
$r(1)$	88,198	121,925	420,114
R-squared	0.503	0.503	0.965
Observations	1981	1981	1981

Notes: Robust standard errors in parentheses, \* significant at 10%, \*\* significant at 5%, \*\*\* significant at 1%. greater than 4 employees.

Table B.3: A model variant without overlap – value added

<i>Estimated parameters: <math>\hat{c}, k</math></i>					
<i>Targeted moments: Share of non-offshoring firms, mean and variance of <math>c = 1 - z</math></i>					
	$\hat{c}$	$k$	$\chi$	$\tilde{c}$	$\text{var}(c)$
Estimates	0.529	1.525	0.384	0.555	0.016
Targets			0.379	0.483	0.154
Difference			0.005	0.072	-0.138
<i>Recovered parameters: <math>\kappa, f, f_e,</math> and <math>\tau L/L^*</math></i>					
	$\kappa$	$f$	$f_e$	$\tau L/L^*$	
Parameters	0.210	593,677	1,000,060	0.971	

Notes: Columns 4-6 in the first and second line of the first panel report the moments obtained from the model (computed using the minimum distance estimates) and the respective moments from the data. The difference between the estimated and the targeted moments are given in the third line.

Table B.4: Model fit: overlap – percentiles

Decile	Overlap		Difference	Decile	Overlap		Difference
	observed	computed			observed	computed	
1	0.308	0	0.308	51	0.83	0.305	0.525
2	0.367	0	0.367	52	0.852	0.338	0.514
3	0.16	0	0.159	53	0.719	0.373	0.346
4	0.409	0	0.409	54	0.571	0.413	0.158
5	0.438	0	0.438	55	0.291	0.457	-0.165
6	0.128	0.001	0.127	56	0.874	0.505	0.369
7	0.945	0.001	0.944	57	0.464	0.557	-0.094
8	0.374	0.001	0.372	58	0.537	0.615	-0.077
9	0.488	0.002	0.486	59	0.801	0.677	0.124
10	0.556	0.002	0.553	60	0.341	0.744	-0.403
11	0.592	0.003	0.589	61	0.89	0.815	0.075
12	0.162	0.003	0.159	62	0.929	0.89	0.039
13	0.716	0.004	0.712	63	0.396	0.968	-0.572
14	0.757	0.005	0.753	64	0.692	0.952	-0.26
15	0.305	0.006	0.299	65	0.251	0.871	-0.62
16	0.983	0.007	0.976	66	0.477	0.79	-0.313
17	0.666	0.008	0.658	67	0.329	0.711	-0.382
18	0.993	0.009	0.985	68	0.912	0.634	0.278
19	0.353	0.01	0.343	69	0.795	0.562	0.233
20	0.83	0.012	0.819	70	0.351	0.495	-0.144
21	0.373	0.013	0.36	71	0.678	0.433	0.245
22	0.405	0.015	0.39	72	0.84	0.376	0.464
23	0.933	0.017	0.916	73	0.142	0.326	-0.184
24	0.559	0.019	0.54	74	0.494	0.281	0.213
25	0.881	0.022	0.86	75	0.275	0.242	0.034
26	0.946	0.024	0.921	76	0.418	0.207	0.211
27	0.283	0.027	0.256	77	0.034	0.177	-0.143
28	0.417	0.03	0.387	78	0.328	0.152	0.176
29	0.273	0.034	0.24	79	0.136	0.129	0.007
30	0.768	0.037	0.73	80	0.106	0.11	-0.005
31	0.51	0.042	0.468	81	0.229	0.094	0.135
32	0.473	0.046	0.427	82	0.332	0.08	0.252
33	0.949	0.051	0.898	83	0.259	0.068	0.19
34	0.532	0.057	0.475	84	0.999	0.058	0.941
35	0.301	0.063	0.239	85	0.38	0.051	0.33
36	0.744	0.069	0.675	86	0.448	0.044	0.404
37	0.856	0.076	0.78	87	0.671	0.039	0.632
38	0.704	0.084	0.62	88	0.826	0.034	0.792
39	0.71	0.093	0.617	89	0.362	0.03	0.332
40	0.861	0.103	0.759	90	0.241	0.026	0.214
41	0.779	0.113	0.666	91	0.262	0.023	0.239
42	0.926	0.125	0.801	92	0.266	0.02	0.245
43	0.945	0.138	0.808	93	0.182	0.018	0.164
44	0.76	0.152	0.608	94	0.651	0.015	0.636
45	0.549	0.168	0.381	95	0.097	0.013	0.084
46	0.496	0.185	0.311	96	0.102	0.01	0.091
47	0.993	0.205	0.788	97	0.236	0.008	0.228
48	0.952	0.226	0.726	98	0.1	0.006	0.094
49	0.68	0.25	0.43	99	0	0.004	-0.004
50	0.606	0.276	0.33	Average	0.536	0.188	0.349