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

Working Paper

Offshoring and Firm Overlap

CESifo Working Paper, No. 6361

Provided in Cooperation with:
Ifo Institute – Leibniz Institute for Economic Research at the University of Munich

Suggested Citation: Capuano, Stella; Egger, Hartmut; Koch, Michael; Schmerer, Hans-Jörg (2017) : Offshoring and Firm Overlap, CESifo Working Paper, No. 6361, Center for Economic Studies and ifo Institute (CESifo), Munich

This Version is available at:
https://hdl.handle.net/10419/155603

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

CESifo Working Papers
www.cesifo.org/wp

Offshoring and Firm Overlap

Stella Capuano
Hartmut Egger
Michael Koch
Hans-Jörg Schmerer

CESIFO WORKING PAPER NO. 6361
CATEGORY 8: TRADE POLICY
FEBRUARY 2017

ISSN 2364-1428

CESifo
Center for Economic Studies & Ifo Institute

CESifo *Working Paper No. 6361*

Offshoring and Firm Overlap

Abstract

We set up a model of offshoring with heterogeneous producers that captures two empirical regularities of German offshoring firms. There is selection of larger, more productive firms into offshoring. However, the selection is not sharp, and offshoring and non-offshoring firms coexist over a wide range of the revenue distribution. An overlap of offshoring and non-offshoring firms emerges in our model because, in contrast to textbook models of trade with heterogeneous producers, we allow firms to differ in two technology parameters thereby decoupling the offshoring status of a firm from its revenues. In an empirical analysis, we employ firm-level data from Germany to estimate key parameters of the model and show that ignoring the overlap lowers the estimated gains from offshoring by more than 50 percent and, at the same time, exaggerates substantially the importance of the extensive margin for explaining the evolution of German offshoring over the last 25 years.

JEL-Codes: F120, F140, L110.

Keywords: offshoring, heterogeneous firms, firm overlap, quantitative trade model, extensive and intensive margins of offshoring.

Stella Capuano
Institute for Employment Research IAB
Nuremberg / Germany
stella.capuano@iab.de

*Hartmut Egger**
University of Bayreuth
Department of Economics
Universitätsstr. 30
Germany – 95447 Bayreuth
hartmut.egger@uni-bayreuth.de

Michael Koch
University of Bayreuth
Bayreuth / Germany
michael.koch@uni-bayreuth.de

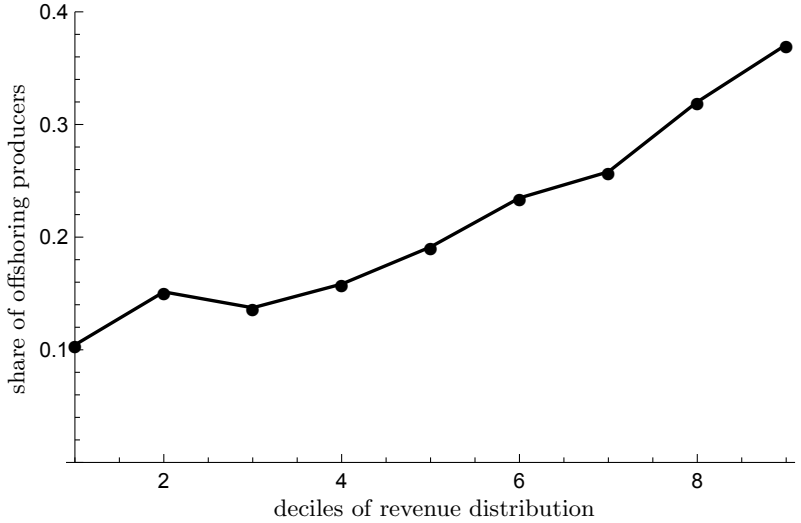
Hans-Jörg Schmerer
Fern Universität Hagen
Hagen / Germany
Hans-Joerg.Schmerer@fernuni-hagen.de

*corresponding author

February 15, 2017
We thank Holger Breinlich, Swati Dhingra, Benedikt Heid, Mario Larch, John Morrow, Peter Neary, Veronica Rappoport, and Esteban Rossi-Hansberg for encouraging discussion on earlier drafts of this manuscript. We are grateful to participants at the Annual Meetings of the European Economic Association, the European Trade Study Group, the International Association for Applied Econometrics, and the German Economic Association, the Danish Trade Workshop, the Dynamics, Economic Growth, and International Trade Conference, the World Congress of the Econometric Society, and the TRISTAN Workshop at the University of Bayreuth as well as research seminars at the Chinese University of Hong Kong, London School of Economics, Shanghai University of Finance and Economics, University of Aarhus, University of Bergen, University of Bielefeld, University of Duesseldorf, University of Duisburg-Essen, University of Essex, University of Hagen, University of Heidelberg, University of Trier, University of Tuebingen, University of Uppsala, University of Wuerzburg, and University of Xiamen for helpful comments.

1 Introduction

Offshoring and its effects on domestic labor markets have played a prominent role in academic research and the public debate over the last two decades. In recent years, attention in the literature has shifted towards understanding the specific nature of firms that choose to offshore. Relying on models of heterogeneous firms, trade economists have pointed out that similar to exporters, offshoring firms are larger, more productive, and make higher profits than their non-offshoring competitors (see Antràs and Helpman, 2004; Antràs et al., 2006; Egger et al., 2015). Although, grosso modo, this pattern is consistent with the data (cf. Bernard et al., 2012; Hummels et al., 2014; Moser et al., 2015), existing theoretical work misses the empirical fact that offshoring and non-offshoring firms coexist over a wide range of the revenue distribution, as illustrated for German producers by Figure 1.[1]



Source: IAB Establishment Panel, covering 20,334 establishment observations for Germany in the years 1999, 2001, 2003 from all industries and size categories; Descriptive statistics refer to own computations, using an interval of 100 millentile observations around the displayed decile positions.

Figure 1: Share of offshoring producers

To explain this fact and to shed light on how it changes the conclusions we draw when it comes to the consequences of offshoring are the aim of this paper. For this purpose, we construct a theoretical model that captures two features, which are characteristic for the empirical pattern of offshoring. *Selection*, because offshoring is more common among producers from

[1] An overlap of offshoring and non-offshoring firms also exists, when using domestic employment or the number of tasks for measuring firm size, or when distinguishing between manufacturing and non-manufacturing producers.

higher quantiles of the revenue distribution; *Overlap*, since there is coexistence of offshoring and non-offshoring producers in the various quantiles of the revenue distribution. After a thorough theoretical analysis, we structurally estimate key parameters of our model, using firm-level data from Germany. Based on these parameter estimates, we then study the nature and extent of the bias in the quantitative welfare effects of offshoring that originates from disregarding the overlap in the data and show how ignoring the overlap affects the relative importance of extensive and intensive margins for explaining observed changes in offshoring.

In the theory section, we set up a two-country model of offshoring, with labor being the only factor of production. The two countries differ in their levels of development and since offshoring is low-cost seeking, it is one directional and leads to production shifting from the more developed source country to the less developed host country. Following Acemoglu and Autor (2011), we model production as the assembly of tasks, with firms differing in the number of tasks performed in the production process. The number of tasks is directly linked to firm productivity, reflecting the idea that more tasks allow for a stronger division of labor in the production of goods. Hence, firm heterogeneity materializes due to differences in the task production process and, in line with our data, this gives a positive link between the number of tasks used and the revenues earned by a firm. Since one source of heterogeneity is not sufficient to model overlap in a setting that features selection into offshoring, we assume that firms also differ in the share of tasks that can be offshored to the low-cost host country.[2] In the tradition of theoretical work building on the Melitz (2003) framework, we model firm heterogeneity as the outcome of a lottery, but acknowledge that firms draw *two* technology parameters: the number of tasks and the share of offshorable tasks.[3]

The interaction of the two technology parameters determines the pattern of offshoring in our model. It is possible that a firm operating a sophisticated technology with many different tasks finds itself in a position with none of its tasks being offshorable, despite its high productivity and large revenues. However, it is also possible that a firm with a simple technology requiring only few tasks can offshore a significant share of these tasks. This provides a source of overlap,

[2]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 25 percent of US jobs as being vulnerable to offshoring according to these criteria.

[3]It is well established that allowing for firm heterogeneity in more than just one dimension helps making the Melitz (2003) model better suited for explaining firm-level evidence in the context of trade. Prominent examples that provide extensions in this direction include Davis and Harrigan (2011), Hallak and Sivadasan (2013), Armenter and Koren (2015), Harrigan and Reshef (2015), and Helpman et al. (2016).

which is rooted in technology and thus exogenous to the firm. To give firms an active role in our model, we assume that offshoring is subject to a fixed cost, and hence the gains from offshoring must be sufficiently high to make it attractive for firms. This makes selection into offshoring a key determinant of overlap.[4] Provided that the fixed costs of offshoring are sufficiently large, offshoring in our model is more attractive ceteris paribus for firms operating a technology with more tasks and thus featuring higher revenues.

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

In the second part of the paper, we combine three datasets from two different sources. The first one is the Establishment Panel of the Institute for Employment Research (IAB), which provides annual survey information on a (varying) sample of up to 16,000 establishments of all branches of the economy and all size categories since the early 1990s. In 1999, 2001, and 2003,

[4]Selection seems important also from an empirical point of view, because the smallest German producers do not offshore.

[5]Relying on the relative effective labor costs when providing intuition for the welfare effects of offshoring acknowledges that trade involves the exchange of final against intermediate goods, so that changes in the relative price of exports and imports do not reflect changes in the terms of trade of consumer goods. We show in the Appendix that a worsening of the factorial terms of trade for the source country is instrumental for the existence of welfare losses from offshoring in the source country.

this dataset also contains information on offshoring activities of German producers, and hence we use firm-level data from these three years in the empirical analysis. As a second source of data input, we rely on the 2006 Employment Survey of the Federal Institute for Vocational Education and Training (BIBB) and the Federal Institute for Occupational Safety and Health (BAuA) to construct a measure of task content for 341 occupations. We finally use the Linked Employer-Employee Database from the Institute for Employment Research (LIAB) to aggregate the task composition at the occupation level to the firm level. This gives a unique dataset for studying offshoring in the context of task production, and we use this dataset to estimate key parameters of our theoretical model, using method of moments.

We estimate the parameters for two model variants: a flexible one, in which we allow for overlap; and a restrictive one, in which we rule out overlap by assumption. We find that the model variant with overlap provides a better fit with the data and show that disregarding the overlap significantly lowers the estimated cost saving from offshoring. This is intuitive, because the model without overlap presumes that all firms that make use of offshoring are high-productivity producers and these firms require a lower cost saving to find offshoring attractive. The discrepancy regarding the estimated cost savings from offshoring generates quantitatively sizable differences in the welfare effects attributed to offshoring by the two models. The model with overlap associates the observed share of offshoring firms with an increase in German GDP per capita of 18.90 percent. The welfare gain attributed to offshoring falls to 8.73 percent and is therefore 53.81 percent (or more than 10 percentage points) lower in the model variant without overlap.

We finally use our quantitative model to decompose the observed increase of German offshoring openness from 18.03 percent in 1990 to 30.26 percent in 2014 into its *intensive* margin – capturing changes in the offshoring activity of incumbent offshoring firms – and its *extensive* margin – capturing changes in the mass of offshoring firms.[6] We show that both margins contributed significantly to the observed increase of German offshoring, with the intensive margin explaining about 45.17 percent of this increase. Disregarding the overlap, the model would attribute only 17.41 percent of the observed increase in German offshoring openness to the intensive margin and therefore considerably exaggerate the role played by the extensive margin. The model with (without) overlap suggests, moreover, that the increase in offshoring openness

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

between 1990 and 2014 has entailed a welfare increase of 4.77 (2.93) percent, which amounts to 12.31 (7.56) percent of the overall increase in German GDP per capita over this period.

Shedding light on the overlap of offshoring and non-offshoring firms, our analysis is most closely related to Armenter and Koren (2015), who have documented an overlap for exporting and non-exporting firms, using US census data. To reconcile the predictions of the Melitz (2003) model regarding the composition of exporters with the data, Armenter and Koren suggest randomizing fixed exporting costs to add an additional source of heterogeneity. Contrasting the thus extended with the original Melitz model, they find that a model with sharp selection into exporting significantly overestimates the role of entry and exit into the export market for the growth of exports. Our analysis differs from Armenter and Koren (2015) in several important ways. First, we document and analyze the overlap of offshoring and non-offshoring firms instead of exporting and non-exporting firms. Second, we root both sources of heterogeneity in the marginal costs of production, which then subsume heterogeneity of firms in all relevant performance measures. This feature allows us to use the toolbox of heterogeneous firms models along the lines of Melitz (2003) for our analysis. Third, we account for dependencies in the distributions of technology parameters, and show that such dependencies are important for describing the overlap in the data. Fourth, we provide a detailed welfare analysis in general equilibrium and show that ignoring the overlap in the data leads to a severe downward bias in the welfare effects predicted by our quantitative trade model.[7]

By studying the effects of offshoring, our model contributes to a large body of literature that includes prominent contributions by Grossman and Rossi-Hansberg (2008), Rodriguez-Clare (2010), and more recently Acemoglu et al. (2015). Thereby, we associate offshoring with a relocation of task production to a low-cost country, as in Grossman and Rossi-Hansberg (2008). However, focussing on the decision of heterogeneous firms to offshore while keeping the share of offshorable tasks constant, we follow Egger et al. (2015) and emphasize a specific adjustment channel, whose quantitative importance has been put forward by recent empirical evidence (cf. Bergin et al., 2011). In addition to adjustments at this extensive firm margin of offshoring, the assumption of a Cobb-Douglas production technology allows for an intensive task margin, which

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

captures changes in the volume of imported tasks at the firm level in response to price changes.

The remainder of the paper is organized as follows. In Section 2 we set up the theoretical model, distinguish important adjustment margins, and study the welfare effects of offshoring in the presence of overlap. In Section 3 we describe the dataset, estimate key model parameters, discuss the goodness of fit of our model, quantify the welfare effects, and show to what extent accounting for the overlap of offshoring and non-offshoring firms in the data affects our results. Section 4 applies our quantitative trade model to decompose the observed increase in German offshoring openness between 1990 and 2014 into its extensive and intensive margin and sheds light on the welfare gains attributable to the increase in offshoring over this period. In Section 5, we study the robustness of our results regarding the composition of host countries and the chosen estimation strategy. The last section concludes with a summary of the most important results.

2 A model of offshoring and firm overlap

2.1 Basic assumptions and intermediate results

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

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

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

The two economies differ in their level of development and are populated by L and L^* units of labor, respectively, where an asterisk refers to the economy with the lower level of development. This is the host country of offshoring, whereas the more advanced economy is the source country of offshoring. Similar to Egger et al. (2015), we assume that the host country lacks the technology

to operate its own firms. This implies that all (industrial) producers are headquartered in the source country and it makes the host country a labor reservoir that is inactive in the absence of offshoring. Firms perform different tasks, which are combined in a Cobb-Douglas technology to produce output $y(\omega)$:

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

where $y(\omega, i)$ denotes the output of task i and $z(\omega)$ is the length of the task interval (or the mass of tasks) performed by firm ω. The technology in Eq. (2) captures in a simple way the gains from labor division, as performing more tasks increases a firm's productivity. Assuming that task output equals labor input, the firm's total variable production costs are given by $C^v(\omega) = \int_0^{z(\omega)} \zeta(i) y(\omega, i) di$, where $\zeta(i)$ is the effective labor cost of task i, which is equal to the domestic wage w if a task is performed at home and equal to the foreign wage w^* multiplied by an iceberg trade cost parameter $\tau > 1$ if the task is performed abroad.

Profit maximization is a three-stage problem. At stage one, (risk-neutral) firms decide on market entry, which involves the investment of f_e units of labor. The investment allows firms to participate in a lottery, in which they draw the length of the task interval z from a common distribution. At stage two, firms decide upon offshoring. This requires the investment of f units of labor and allows them to draw technology parameter s in a second lottery. After the investment, firms can put the share s of their tasks offshore.[8] At stage three, firms hire workers, produce, and sell their output in a monopolistically competitive market, facing consumer demand in Eq. (1). Being a monopolist in their own market, firms consider $x(\omega) = y(\omega)$ and thus the impact of their employment decision on their own price. At the same time, firms are atomistic in the aggregate, and hence take income I and price index P as given. We solve the three-stage problem by backward induction.

At stage three firms make the employment decision for each task at home and abroad. Due to the underlying Cobb-Douglas technology in Eq. (2), profit maximization establishes the result that expenditures are the same for all tasks. The marginal production cost of firm ω is therefore

[8] Two remarks are in order here. First, one could also assume that firms draw s together with z in the first-stage lottery. Provided that s is not revealed prior to the investment of stage two, this would give the same formal structure as our approach. Second, we could also choose the unit cost of offshoring, τ, instead of the share of offshorable tasks, s, as the second variable contributing to firm heterogeneity. We decided against this alternative, because we use changes in the unit offshoring cost parameter in a comparative static analysis to shed light on how firms respond in their offshoring decision to a symmetric cost shock.

given by

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

where $\kappa \equiv \tau w^*/w$ denotes *effective* labor costs in the host relative to the source country of offshoring. Since offshoring has fixed costs, $\kappa < 1$ must hold in order to make it attractive for firms to shift task production abroad, and we can associate $\kappa^{-s(\omega)}$ with the marginal *cost saving effect* of offshoring. Due to an isoelastic demand function, profit maximization at stage three further establishes the well-known result that firms set their prices as a constant markup over their marginal costs: $p(\omega) = c(\omega)\sigma/(\sigma - 1)$. In view of Eq. (1), firm-level revenues, $r(\omega) = p(\omega)y(\omega)$ are then given by $r(\omega) = I[p(\omega)/P]^{1-\sigma}$, and relative revenues of two firms can be expressed as a decreasing function of their marginal cost differential:

$$\frac{r(\omega_1)}{r(\omega_2)} = \left[\frac{c(\omega_1)}{c(\omega_2)}\right]^{1-\sigma}. \tag{4}$$

In view of Eq. (4), we can index revenues r by marginal costs c instead of ω from now on, in the understanding that marginal production costs are firm-specific.

At stage two firms make their offshoring decision. Offshoring requires the investment of $f > 0$ units of labor, which allows firms to participate in a lottery, in which they draw the share of offshorable tasks s. We assume that the distribution of s depends on the realization of z. To be more specific, 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 assume $\Pr_z(s > 0) \equiv \nu_0 + \nu_1 z$, with $\nu_0, \nu_1 \geq 0$ and $\nu \equiv \nu_0 + \nu_1 \in (0, 1]$. Furthermore, for firms with some offshorable tasks, the share of tasks that can be put offshore, s, is uniformly distributed over the interval $(0, 1]$. Hence, for a firm with task length z, the ex ante expected value of s is given by $\mathbb{E}_z[s] = (\nu_0 + \nu_1 z)/2$. The expected *relative* revenue gain from offshoring depends on the cost saving under all possible realizations of s. For $\nu_1 > 0$, it is larger for firms with a better z-draw: $\mathbb{E}_z\left[\kappa^{s(1-\sigma)}\right] = \Pr_z(s > 0)\int_0^1 \kappa^{s(1-\sigma)}ds$, with $d\Pr_z(s > 0)/dz > 0$. In *absolute* terms, there is a second advantage that renders offshoring more attractive for firms with a better z-draw. They make higher revenues at any possible realization of s, according to Eqs. (3) and (4), and hence

can more easily cover the fixed cost of offshoring.[9]

Being risk-neutral, firms will make the offshoring investment only if its expected return is sufficiently high, and since the expected return is higher ceteris paribus for firms that have drawn a larger value of z in the lottery, our model establishes for a sufficiently high fixed cost parameter f selection of high-productive firms into offshoring. For the moment, we simply assume selection, whereas in Section 2.3 we characterize the parameter domain that supports selection in our model. Accounting for Eqs. (3) and (4), the expected profit gain from offshoring of a firm with task length z can be expressed as $\Pr_z(s>0)(1-z)^{1-\sigma}r(w)[\int_0^1 \kappa^{s(1-\sigma)}ds-1]/\sigma-fw$, where $r(w)$ is the revenue of the least productive firm with $z=0$ and $c=w$, which is a firm that does not offshore (see below). The marginal offshoring firm with task length $\hat{z}$, which is the firm that is indifferent between making and not making investment f, is therefore characterized by the following condition

$$\sigma f w = \left(\nu - \nu_1 \frac{\hat{c}}{w}\right)\left(\frac{\hat{c}}{w}\right)^{1-\sigma} r(w)\left[\frac{\kappa^{1-\sigma}-1}{(1-\sigma)\ln\kappa}-1\right]. \tag{5}$$

where $\hat{c} \equiv (1-\hat{z})w$.

At stage one, firms decide on firm entry. To enter the source country, they must make an initial investment of $f_e > 0$ units of labor. This investment gives them a single draw of task length z from a common distribution function. For tractability reasons, we assume that z is Pareto distributed over the unit interval with a probability density function $g_z(z) = k(1-z)^{k-1}$, $k>0$. We consider a static model and, following Ghironi and Melitz (2005), abstract from fixed costs of production, so that all firms participating in the technology lottery start production, irrespective of their z-draw. We do not allow for selection into production, because our dataset covers many small producers, which employ only few domestic workers. Free entry requires that firms make zero profits in expectation, and hence that aggregate operating profits, i.e. total

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

revenues R divided by σ, are equal to economy-wide expenditures for fixed costs, $M\hat{c}^k f + Mf_e$, where M is the mass of firms producing distinct varieties ω. The solution to the firms' problem at stage one gives the mass of firms entering the z-lottery, which is determined in general equilibrium and discussed in Section 2.3. To solve for the general equilibrium outcome, we first need to understand how the distributions of the two technology parameters z and s determine the distribution of marginal costs (and thus revenues) in our setting.

2.2 The distribution of marginal costs

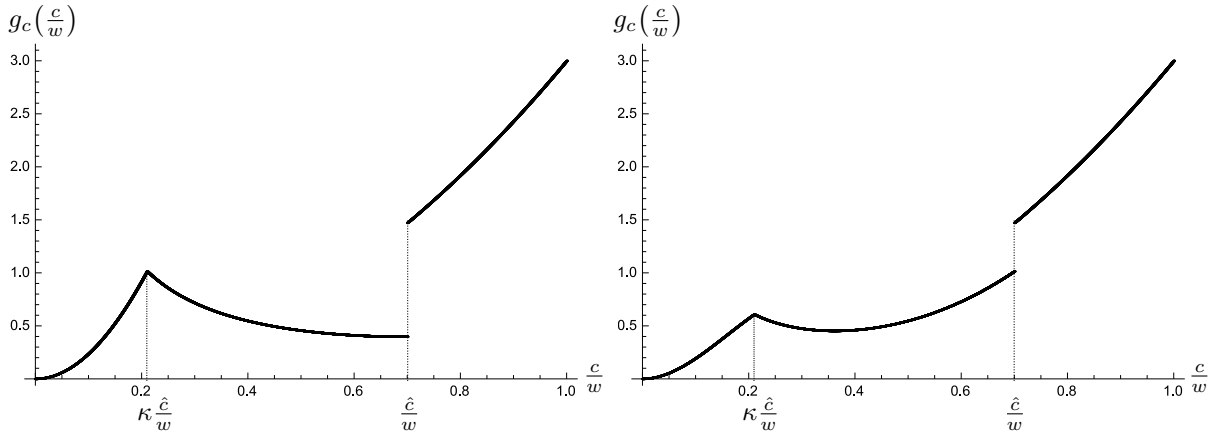
Even though our model features two forms of firm heterogeneity, we can conclude from Eq. (3) that their combined effect on firm-level performance measures is captured by a single variable: the marginal cost of production. This implies that we can learn about the distribution of firms in their various performance measures, when we understand how the distributions of the two technology parameter z and s map into the distribution of marginal costs c. The marginal cost of non-offshoring firms is given by $c = (1 - z)w$, according to Eq. (3). Non-offshoring firms are either low-productivity producers with task length $z \leq \hat{z}$ or they are high productivity producers with task length $z \geq \hat{z}$ and no offshorable task. Due to the inverse link between c and z, there is no difference between ranking non-offshoring firms by their task length or the marginal costs – with the ordering of firms flipped – and for these firms we can therefore infer the distribution of marginal costs c from the distribution of task length z and the insights that a z-specific share of firms, 1-$\Pr_z(s > 0)$, has not a single offshorable task.

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

costs c/w:

$$g_c\left(\frac{c}{w}\right) = \begin{cases} \left(1-\nu+\nu_1\frac{c}{w}\right)k\left(\frac{c}{w}\right)^{k-1} - \frac{1}{\ln\kappa}\left\{\nu\left(\frac{c}{w}\right)^{k-1}\left[\left(\frac{1}{\kappa}\right)^{k}-1\right] - \nu_1\frac{k(c/w)^k}{k+1}\left[\left(\frac{1}{\kappa}\right)^{k+1}-1\right]\right\} & \text{if} \quad \frac{c}{w} \leq \kappa\frac{\hat{c}}{w} \\ \left(1-\nu+\nu_1\frac{c}{w}\right)k\left(\frac{c}{w}\right)^{k-1} - \frac{1}{\ln\kappa}\left\{\nu\left(\frac{c}{w}\right)^{k-1}\left[\left(\frac{\hat{c}/w}{c/w}\right)^{k}-1\right] - \nu_1\frac{k(c/w)^k}{k+1}\left[\left(\frac{\hat{c}/w}{c/w}\right)^{k+1}-1\right]\right\} & \text{if} \quad \frac{c}{w} \in \left(\kappa\frac{\hat{c}}{w}, \frac{\hat{c}}{w}\right] \\ k\left(\frac{c}{w}\right)^{k-1} & \text{if} \quad \frac{c}{w} > \frac{\hat{c}}{w} \end{cases} \cdot \tag{6}$$

The probability density function of c/w is illustrated for two different sets of parameters in Figure 2.



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

Figure 2: The probability density function $g_c\left(\frac{c}{w}\right)$

As we can see from Eq. (6) and Figure 2 the pdf of (normalized) marginal costs, $g_c\left(\frac{c}{w}\right)$, has support on the unit interval and features a discontinuity at $\hat{c}/w$. This is because for firms with task length $z \geq \hat{z}$ investment into offshoring is attractive, and a subset of these firms detects to have at least some offshorable tasks and thus starts offshoring. Since offshoring firms experience a marginal cost saving and are thus shifted to a lower c/w and since the fraction of firms that is affected by this cost saving is discrete for any $z > 0$, selection into offshoring generates a discontinuity of the pdf at $\hat{c}/w$ in Figure 2. The kink of the pdf function at $\kappa\hat{c}/w$ is also rooted in the selection of high-productivity firms into offshoring. More specifically, since firms with $z < \hat{z}$ refuse to make the fixed cost investment for learning about the offshorability of

their tasks, none of these firms is shifted towards lower marginal costs. This imposes a binding (selection) constraint on the number of firms that can be located at the marginal cost interval $\left(\kappa\frac{\hat{c}}{w}, \frac{\hat{c}}{w}\right]$. For (normalized) marginal costs $c/w < \kappa\hat{c}/w$ the selection constraint is not binding, because the maximum possible cost saving from offshoring when shifting all tasks abroad is given by κ, and hence a firm with task length $z < \hat{z}$ could not be shifted to a (normalized) marginal cost lower than $\kappa\hat{c}/w$ even if it would make the investment into offshoring despite an expected profit loss.

2.3 The general equilibrium

To solve for the general equilibrium, we choose source country labor as numéraire and set $w = 1$. As shown in the Appendix, using Eq. (6), we can express economy-wide revenues as follows:

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

where $r(1)$ is the revenue of the least productive producer if $w = 1$ and $k > 2(\sigma-1)$ is assumed to ensure a finite positive value of both the mean and the variance of revenues (cf. Helpman et al., 2004). As outlined above, free entry establishes $R = M\sigma\left(f_e + \hat{c}^k f\right)$. Together with Eqs. (5) and (7), this gives a relationship between the marginal cost of the offshoring firm that is indifferent between making and not making investment f, $\hat{c}$, and the effective wage differential between the host and the source country of offshoring, κ, which we call *offshoring indifference condition* (OC):

$$\Gamma_1(\hat{c},\kappa) \equiv \frac{\hat{c}^{\sigma-1}}{\nu-\nu_1\hat{c}}\frac{k}{k-\sigma+1} + \left\{\frac{\hat{c}^k}{\nu-\nu_1\hat{c}}\left[\frac{(\sigma-1)\nu}{k-\sigma+1} - \frac{(\sigma-2)\nu_1\hat{c}}{k-\sigma+2}\right] - \frac{f_e}{f}\right\}\left[\frac{\kappa^{1-\sigma}-1}{(1-\sigma)\ln\kappa} - 1\right] = 0. \qquad (8)$$

As formally shown in the Appendix, $\Gamma_1(\cdot) = 0$ establishes a negative link between $\hat{c}$ and κ. The larger the relative effective labor costs in the host country are, the smaller is the cost saving effect of offshoring and the more productive the marginal firm that makes investment f must be in order to avoid in expectation losses from this investment. Intuitively, if the cost saving from offshoring vanishes due to $\kappa = 1$, all firms prefer domestic production, resulting in $\hat{c} = 0$.

In contrast, $\hat{c}$ reaches a maximum at low levels of κ.

A second link between $\hat{c}$ and κ can be determined, when noting from above that free entry into the technology lottery at stage one implies that all (disposable) income accrues to workers, $I = L + w^*L^*$. Since global income is equal to total consumption expenditures, we have $R = L + w^*L^*$. Furthermore, constant markup pricing establishes the well-known result that variable production costs are a constant fraction $(\sigma - 1)/\sigma$ of a firm's revenues, with part of these costs accruing to imported tasks. As formally shown in the Appendix, the wage bill for workers in the host country of offshoring can thus be expressed as a function of aggregate revenues according to

$$w^*L^* = R\frac{\sigma-1}{\sigma}\frac{\hat{c}^{k-\sigma+1}\left(\nu-\nu_1\hat{c}^{\frac{k-\sigma+1}{k-\sigma+2}}\right)\frac{1+\kappa^{1-\sigma}[(1-\sigma)\ln\kappa-1]}{[(1-\sigma)\ln\kappa]^2}}{1+\hat{c}^{k-\sigma+1}\left(\nu-\nu_1\hat{c}^{\frac{k-\sigma+1}{k-\sigma+2}}\right)\left(\frac{\kappa^{1-\sigma}-1}{(1-\sigma)\ln\kappa}-1\right)}. \tag{9}$$

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

$$\Gamma_2(\kappa,\hat{c}) \equiv \kappa\left\{\frac{\sigma}{\sigma-1}\frac{1+\hat{c}^{k-\sigma+1}\left(\nu-\nu_1\hat{c}^{\frac{k-\sigma+1}{k-\sigma+2}}\right)\left(\frac{\kappa^{1-\sigma}-1}{(1-\sigma)\ln\kappa}-1\right)}{\hat{c}^{k-\sigma+1}\left(\nu-\nu_1\hat{c}^{\frac{k-\sigma+1}{k-\sigma+2}}\right)\frac{1+\kappa^{1-\sigma}[(1-\sigma)\ln\kappa-1]}{[(1-\sigma)\ln\kappa]^2}}-1\right\}-\frac{\tau L}{L^*}=0 \tag{10}$$

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

The equilibrium values of $\hat{c}$ and κ are jointly determined by the offshoring indifference condition and the labor market constraint. Thereby, our model features a unique interior equilibrium if offshoring cost parameters τ and f are sufficiently high.[10] The impact of changes in the two offshoring cost parameters is illustrated in Figure 3. A higher variable offshoring cost parameter

[10]The critical levels of τ and f depend – among other model parameters – on the levels of ν_0 and ν_1. In the knife-edge case of $\nu_0 = 0$, a unique interior equilibrium is guaranteed for any combination of τ and f.

τ implies for a given volume of offshoring that more foreign workers must be employed in order to provide the required amount of tasks for production in the source country. Therefore, the effective cost for employing foreign relative to domestic labor, κ, must increase to restore labor market clearing. This effect is captured by a counter-clockwise rotation of locus LC in Figure 3, which makes an interior solution with intersection of OC and LC at $\hat{c} < 1$ and $\kappa < 1$ more likely. A higher offshoring fixed cost parameter makes offshoring less attractive ceteris paribus and therefore lowers the cutoff cost level characterizing the firm that is indifferent between making and not making the investment of f. This effect is captured by a clockwise rotation of locus OC in Figure 3, which also makes the existence of an interior equilibrium more likely.

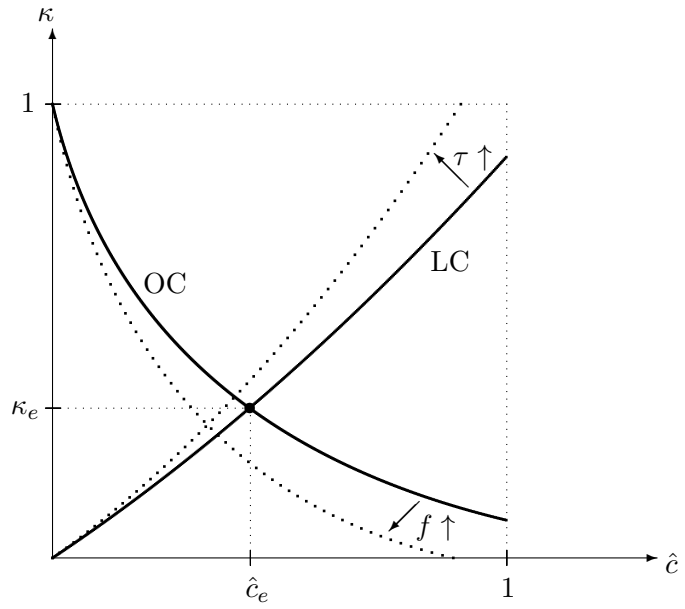


Figure 3: Equilibrium values of $\hat{c}$ and κ

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

present when the variable offshoring parameter increases, it is counteracted and dominated by the initial increase in τ, so that the effective wage differential increases.

2.4 Offshoring margins and welfare

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

$$\chi(c) = \begin{cases} 1 - \left[1 - \frac{1}{\ln \kappa} \frac{\nu\left[\left(\frac{1}{\kappa}\right)^k - 1\right] - \nu_1 \frac{kc}{k+1}\left[\left(\frac{1}{\kappa}\right)^{k+1} - 1\right]}{(1-\nu+\nu_1 c)k}\right]^{-1} & \text{if} \quad c \leq \kappa\hat{c} \\ 1 - \left[1 - \frac{1}{\ln \kappa} \frac{\nu\left[\left(\frac{\hat{c}}{c}\right)^k - 1\right] - \nu_1 \frac{kc}{k+1}\left[\left(\frac{\hat{c}}{c}\right)^{k+1} - 1\right]}{(1-\nu+\nu_1 c)k}\right]^{-1} & \text{if} \quad c \in (\kappa\hat{c}, \hat{c}] \\ 0 & \text{if} \quad c > \hat{c} \end{cases} \tag{11}$$

according to Eq. (6). It is easily confirmed that $\chi'(c) < 0$ holds for all $c < \hat{c}$, implying that the share of offshoring firms decreases in c. The economy-wide share of offshoring firms is then given by the frequency-weighted mean of $\chi(c)$ and amounts to

$$\chi = \hat{c}^k \left[\nu - \nu_1 \frac{k}{k+1}\hat{c}\right]. \tag{12}$$

From Eq. (12) we see that the share of offshoring firms, χ, increases in the cutoff level of marginal costs $\hat{c}$: $d\chi/d\hat{c} = k\hat{c}^{k-1}(\nu - \nu_1\hat{c}) > 0$. Since we know from Figure 3 that $d\hat{c}/df < 0$ and $d\hat{c}/d\tau < 0$, we can thus conclude that a decline in either offshoring cost parameter increases the share of offshoring firms and thus raises offshoring along the extensive margin.

To study adjustments of offshoring along the intensive margin, we can note that total task

expenditures of offshoring firms are given by $[(\sigma-1)/\sigma]R[1-R^d/R]$, with

$$\frac{R^d}{R}=\frac{1-\hat{c}^{k-\sigma+1}\left(\nu-\nu_1\hat{c}^{\frac{k-\sigma+1}{k-\sigma+2}}\right)}{1+\hat{c}^{k-\sigma+1}\left(\nu-\nu_1\hat{c}^{\frac{k-\sigma+1}{k-\sigma+2}}\right)\left(\frac{\kappa^{1-\sigma}-1}{(1-\sigma)\ln\kappa}-1\right)}, \tag{13}$$

being the fraction of aggregate revenues accruing to non-offshoring producers. In view of Eq. (9), we can thus write the expenditure share of offshoring firms for imported tasks as follows

$$\rho\equiv\frac{w^*L^*}{[(\sigma-1)/\sigma]R[1-R^d/R]}=\frac{\kappa^{1-\sigma}}{\kappa^{1-\sigma}-1}-\frac{1}{(1-\sigma)\ln\kappa}, \tag{14}$$

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

A distinction between the extensive and intensive margin of offshoring is important for understanding the welfare implications of offshoring in the source country.[11] Since preferences are homothetic, we can use the representative consumer in a normative interpretation and consider per-capita labor income ($\widehat{=}$ GDP per capita) as our preferred welfare measure. In view of $w=1$, we can thus express source country welfare as the inverse of the consumer price index: $W=P^{-1}$. To determine the consumer price index, we can start from the observation that revenues are the product of prices and output. Therefore, accounting for Eq. (1) and our previous insight that global consumption expenditure is equal to total source and host country labor income $L+w^*L^*$,

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

revenues of the least productive firm can be expressed as

$$r(1) = \frac{L + w^* L^*}{P^{1-\sigma}} \left(\frac{\sigma - 1}{\sigma} \right)^{\sigma - 1}. \tag{15}$$

A second expression for the revenues of the least productive producer can be found when combining the indifference condition of the marginal offshoring firm in Eq. (5) with the offshoring indifference condition $\Gamma_1(\cdot) = 0$:

$$r(1) = \sigma f \left[\frac{f_e}{f} - \frac{\hat{c}^k}{\nu - \nu_1 \hat{c}} \left(\nu \frac{\sigma - 1}{k - \sigma + 1} - \nu_1 \hat{c} \frac{\sigma - 2}{k - \sigma + 2} \right) \right] \frac{k - \sigma + 1}{k}. \tag{16}$$

The two Eqs. (15) and (16) jointly determine price index P and thus source country welfare

$$W = \left\{ \frac{L + \kappa L^*/\tau}{\sigma f} \left[\frac{f_e}{f} - \frac{\hat{c}^k}{\nu - \nu_1 \hat{c}} \left(\nu \frac{\sigma - 1}{k - \sigma + 1} - \nu_1 \hat{c} \frac{\sigma - 2}{k - \sigma + 2} \right) \right]^{-1} \frac{k}{k - \sigma + 1} \right\}^{\frac{1}{\sigma - 1}} \frac{\sigma - 1}{\sigma}. \tag{17}$$

A decline in τ induces an expansion of offshoring along both the intensive and extensive margin and therefore raises foreign labor demand. Whereas this leads to higher foreign wages, the increase in the foreign wage rate is not strong enough to dominate the initial decline in the variable offshoring cost. As a consequence, the effective foreign labor cost decreases, reflecting an appreciation of domestic relative to foreign labor and thus an improvement in the source country's (double) factorial terms of trade, with positive welfare consequences. Things are different if the fixed cost of offshoring decreases, because the expansion of offshoring along the extensive margin not only raises foreign wages but also the relative effective cost of employing workers in the host country. This induces a decline of offshoring along the intensive margin and worsens the (double) factorial terms of trade of the source country. The depreciation of domestic relative to foreign labor may be strong enough to dominate the source country's direct welfare gain from a lower offshoring fixed cost. In the Appendix, we provide a formal discussion of these effects and illustrate the possibility of welfare losses for the source country from a lower fixed offshoring cost by means of a numerical example.[12]

Welfare in the source country and the relative importance of the extensive and the intensive margin of offshoring are the two main targets of the empirical analysis conducted in Sections

[12] Welfare losses in the source country do not go hand in hand with global welfare losses, because the host country benefits from higher labor demand.

3 and 4. There, we use the formal structure of our model as guidance to estimate the main parameters of this model and to analyze the aptitude of our model to capture important features of the data. Furthermore, we will shed light on how important acknowledging the observed overlap is for quantifying the welfare effects of offshoring and for assessing the relative importance of the two margins of offshoring.

3 An empirical analysis

To make our model accessible to a structural empirical analysis, we collect data from two different sources and estimate key parameters of our model using a structural approach. In the next two subsections, we describe the data and outline the empirical methodology. The parameter estimates from our empirical analysis are reported in Subsection 3.3. In Subsection 3.4, we discuss the fit of our quantitative model with the data, and in Subsection 3.5 we use the parameter estimates to quantify the welfare effects of offshoring for Germany.

3.1 Data sources and descriptives

We use data on revenues and offshoring of German firms from the Establishment Panel of the Institute of Employment Research (IAB) in Nuremberg. This database reports detailed establishment information from employer surveys at an annual basis since 1993.[13] However, information on the offshoring activity of German firms is only available in the years 1999, 2001, and 2003. Following Moser et al. (2015), we associate offshoring with the purchase of intermediates or other inputs from abroad in the previous business year. Dropping firms that lack either offshoring or revenue information, gives us a sample of 12,250 different firms and 20,341 firm observations for the three years. We do not exploit the time-series variation in the data and rely on the observational information to construct a cross section of German firms for estimating the parameters of the static model outlined in Section 2.

Building on the idea of task offshoring it is important for our structural approach to gather information on the number of tasks performed in German firms. To construct this data, we rely on the BIBB/BAuA Employment Survey, which provides information on workplace characteris-

[13]The IAB Establishment Panel provides plant-level information, which unfortunately cannot be aggregated to firm-level information with the available data. Although there are obvious differences between establishments and firms, we follow previous work and in the interest of readability use the two terms interchangeably throughout our empirical analysis.

tics, including the tasks performed and the occupations held by respondents, for a representative sample of German employees with a working time of more than 10 hours per week (see Rohrbach-Schmidt, 2009, for a detailed description). Interviews have been conducted six times since 1979. Given the temporal proximity to the offshoring events in the IAB Establishment Panel, we use information on tasks and occupations from the 2006 survey, which covers 20,000 employees. In this survey, we can distinguish 28 different tasks (activities), which are listed in the Appendix. Depending on the content, interviewees can answer the 28 questions on whether they perform a certain task either with *often/sometimes/never* or with *yes/no*. We map a task to an occupation if at least 2/3 of the respondents in an occupation declared to perform that task to some extent – i.e. if their answer was *often*, *sometimes*, or *yes*. We distinguish 341 different occupation units (*Berufsordnungen*), using the classification of the Federal Employment Agency, KldB 1988, and collect for each of these occupations the number of tasks performed, according to the responses of the interviewees.[14]

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

From the first row of Table 1, we see that 23 percent of German firms offshore. This figure is somewhat higher than the share of offshoring firms reported by Moser et al. (2015). The reason

[14]By using information from the BIBB/BAuA Employment Survey to learn about the task content of occupations, we follow Spitz-Oener (2006) and Becker et al. (2013).

Table 1: Descriptive Statistics

	Mean	Median	STD.	Min.	Max.
Offshoring	0.23	0.00	0.42	0.00	1.00
Task interval	0.30	0.29	0.14	0.04	0.89
Tasks per occupation	5.24	5.23	1.67	0.19	16.00
Workforce with soc. sec.	12.26	3.00	88.83	0.00	42,291
Revenues in million Euro	2.46	0.35	35.56	0.00	12,055

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

for this difference is that Moser et al. (2015) define offshoring by a qualitative increase of a firm's share of foreign intermediates in two consecutive periods with offshoring information, which is a more restrictive definition of offshoring than the one we use in our cross section. The second row tells us that the firms in our dataset are nicely distributed over the unit task interval. In the third row we see that the typical occupation in a German firm is a multi-task entity and that the employees perform more than five tasks in an average firm. Firms with less than one task per occupation are firms that employ laundry workers and pressers to whom we cannot assign a single task (see above). Since we consider all firms, for which we have the required offshoring, revenue, and task information, our dataset features large differences of firms in both the size of workforce and the size of revenues.[15]

3.2 Estimation strategy

For a quantification of our model, we have to gauge information on the six structural parameters ν_0, ν_1, σ, k, $\hat{c}$, and κ. Since the six parameters jointly determine the observables in our dataset, we cannot estimate them independently, using linear specifications. Rather, we have to solve a system of equations in order to estimate the six parameter values. This is a difficult task, and we therefore reduce the dimensionality of the estimation problem by making use of structural relationships from our model. First of all, our model produces an inverse relationship between

[15]Since the aggregation of occupations to firms using workforce information from LIAB is also possible if the employment status is not subject to social security payments, our dataset includes small firms that do not have a single employee for whom they have a legal obligation to make such payments. Revenues are positive for all firms and a value of zero for the minimum is simply the result of rounding.

the quantile position of firms in the revenue distribution, q, and their quantile position in the marginal cost distribution, $1-q$. This allows us to link the cutoff of marginal costs separating firms that make an investment into offshoring from firms that do not make this investment, $\hat{c}$, to the quantile position of the most productive non-offshoring firm, $\hat{q}$: $\hat{c} = (1-\hat{q})^{1/k}$. Since $\hat{q}$ can be observed in the data, we can express $\hat{c}$ as a function of k. Moreover, we can make use of Eq. (12) and the observed share of offshoring firms, $\widehat{\chi}$, to determine the $(\hat{c}, k)$-combinations that are consistent with feasible realizations of ν_0 and ν_1. Employing $\hat{c} = (1-\hat{q})^{1/k}$ from above, then gives k as a function of parameter tuple (ν_0, ν_1). Finally, we can use the observed revenue ratio $\widehat{R_d/R}$ in Eq. (13) to determine the $(\hat{c}, k, \kappa)$-combinations that are consistent with feasible realizations of the three parameters ν_0, ν_1, and σ. Accounting for the solutions of $\hat{c}$ and k from above, we can solve for κ as a function of the parameter triple (ν_0, ν_1, σ). Following this reasoning, we can reduce the estimation problem to one, in which we simultaneously determine the three remaining parameters ν_0, ν_1, and σ, while recovering parameters $\hat{c}$, k and κ from the structural relationships imposed by our model.

Estimation of ν_0, ν_1, and σ: We use a minimum distance Method-of-Moments (MM) estimator outlined in Ferguson (1958) and Cameron and Trivedi (2005). This estimator is similar to other MM applications and builds on the idea to specify a vector of n_m observed population moments, $\mathbf{m}$, which is linked to a vector of n_x structural parameters of the model, $\mathbf{x}$, according to $\mathbf{m} = \mu(\mathbf{x})$, where $\mu(\mathbf{x})$ is a $n_m \times 1$ vector function. If the number of moments, n_m, is larger than the number of structural parameters, n_x, we can estimate the structural parameters $\mathbf{x}$ by minimizing the weighted squared distance between observed moments $\mathbf{m}$ and computed moments $\mu(\mathbf{x})$, subject to a vector of constraints, **Cons** that are imposed by the theoretical model:

$$\widehat{\mathbf{x}}_{MD} = \text{argmin}_{\mathbf{x}} \ \left(\widehat{\mathbf{m}} - \mu(\mathbf{x})\right)' \mathbf{W} \left(\widehat{\mathbf{m}} - \mu(\mathbf{x})\right), \quad s.t. \quad \textbf{Cons}, \tag{18}$$

where $\mathbf{W}$ is a $n_m \times n_m$ positive-semidefinite weighting matrix and a hat indicates observed or estimated variables. The specific assumption of the MM estimator considered here is that $\mathbf{m}$ is a vector of reduced-form parameters, whose estimates $\widehat{\mathbf{m}}$ are the means of subsets of observations. As weighting matrix $\mathbf{W}$, we use a diagonal matrix based on the inverse variances of the observations used to construct the reduced-form parameter estimates. This puts higher

weight on more precisely measured moments and is the optimal weighting matrix for given reduced-form estimates $\widehat{\mathbf{m}}$ (cf. Cameron and Trivedi, 2005).[16]

We consider four moments in Eq. (18). The first one is the variance of marginal costs, which in turn are inversely related to the number of tasks performed in an establishment. Since marginal costs are not directly observable and information on the tasks performed in an occupation is only available for the German workforce, we restrict our analysis to non-offshoring establishments. For these producers, we compute the variance of marginal costs from information of the task content of production, according to $c = 1 - z$. The variance of marginal costs of non-offshoring establishments gives the first data moment for our estimation: $m_1 = 0.02$. Using Eq. (6), the theoretical counterpart of the variance of marginal costs of non-offshoring producers is:

$$\mu_1(\nu_0, \nu_1, k) = \frac{k}{k+2} \frac{1 - \hat{c}^{k+2}\left(\nu - \nu_1 \frac{k+2}{k+3}\hat{c}\right)}{1 - \hat{c}^{k}\left(\nu - \nu_1 \frac{k}{k+1}\hat{c}\right)} - \left[\frac{k}{k+1} \frac{1 - \hat{c}^{k+1}\left(\nu - \nu_1 \frac{k+1}{k+2}\hat{c}\right)}{1 - \hat{c}^{k}\left(\nu - \nu_1 \frac{k}{k+1}\hat{c}\right)}\right]^2, \tag{19}$$

with $\nu = \nu_0 + \nu_1$ and $\hat{c} = (1 - \hat{q})^{1/k}$.

To construct the second moment, we make use of the insight from Eq. (4) that the revenue of any two establishments is a power function of their relative marginal costs. Hence, we can express the *average* log revenue of non-offshoring establishments relative to the log revenue of a non-offshoring producer with marginal cost c_1, according to $\mathbb{E}(\ln r|\text{not offshoring}) = \ln r(c_1) + (1-\sigma)[\mathbb{E}(\ln c|\text{not offshoring}) - \ln c_1]$. For non-offshoring establishments we can infer the level of marginal costs from the observed usage of tasks, according to $c = 1-z$ (see above). This allows us to compute average log marginal costs of non-offshoring establishments: $\mathbb{E}(\ln c|\text{not offshoring}) = -0.35$. A natural candidate for c_1 is the highest observed marginal cost in our dataset, which amounts to 0.96. There are 118 establishments sharing these costs. To guard ourselves against outliers, we choose the firm at the 25th percentile of the revenue distribution among those 118 establishments as an anchor for our estimation and set $\ln c_1 = -0.04$, $\ln r(c_1) = 11.15$.[17] Putting

[16] In a robustness analysis presented in Section 5 we analyze how our results change when relying on an estimate of the inverse variance-covariance matrix of moment conditions for constructing $\mathbf{W}$, as it is common in GMM applications.

[17] Our results are robust to the specific percentile position chosen for comparison.

together, we compute the mean of log revenues of non-offshoring establishments according to

$$\mu_2(\sigma) = 11.15 + 0.32(\sigma - 1). \tag{20}$$

The empirical counterpart to this moment can be directly observed in the data and amounts to $m_2 = 12.75$.[18]

We use the two remaining moments to capture the overlap by targeting the share of offshoring establishments at the first and the ninth deciles of the revenue distribution. To construct these moments, we make use of Eq. (6) and compute the marginal cost corresponding to a certain quantile position in the revenue distribution according to $c_q = (1-q)^{1/k}$. With the marginal cost level at hand, the share of offshoring establishments at quantile position q is then given by $\chi(c_q)$, according to Eq. (11). This gives $\mu_j(\nu_0, \nu_1, k, \kappa; q)$ for moment $j = 3, 4$, where q refers to the quantile position of the establishment with marginal cost c_q in the revenue distribution of *all* offshoring and non-offshoring establishments. When determining the observed counterparts for moments 3 and 4 – denoted by m_3 and m_4, respectively – we face the problem that each decile position covers only a small fraction of establishments, which potentially makes the data moments vulnerable to outliers and, at the same time, challenges the idea that inverse variances provide suitable weights for these moment conditions. To avoid such problems, we construct the observed share of offshoring establishments at a decile position by averaging over an interval of 100 millentile observations around the respective decile position.[19] This gives $m_3 = 0.10$ and $m_4 = 0.37$ for the share of offshoring firms at the first and ninth decile, respectively.

Finally, in order for the estimated parameters to satisfy the assumptions of the model, we

[18] We could also compute mean and variance of revenues as function of the marginal producer, using the revenue distribution in our model. This would give more complicated expressions but, at the same time, would have at least two attractive features. On the one hand, the thus determined expressions would not require to use information on marginal costs from the data. On the other hand, targeting the variance of revenues would enforce constraint $k > 2(\sigma - 1)$, and hence make adding this condition as separate parameter constraint obsolete (see below). Unfortunately, when following this approach we are not able to find a unique numeric solution for minimization problem (18) in the model variant without overlap, and hence we have decided against this alternative.

[19] To be more specific, we associate the share of offshoring establishments at the first and ninth decile with the average shares over millentiles $\{45, ..., 100, ..., 144\}$ and millentiles $\{845, ..., 900, ..., 944\}$, respectively.

specify the vector of constraints **Cons** in problem (18) as follows:

$$\mathbf{Cons} = \begin{bmatrix} \sigma > 1 \\ k > 2(\sigma - 1) \\ \nu_0, \nu_1 \geq 0 \\ \nu_0 + \nu_1 \leq 1 \end{bmatrix}.$$

Implementation of the estimation strategy: Since the moment conditions outlined above are highly nonlinear functions of the parameters of our model, we cannot solve the minimization problem (18) analytically. Therefore, we choose a numerical approach, consider a discrete parameter space with fine grid, and compute the theory moments for all possible combinations for ν_0 ν_1, σ and the corresponding values of $\hat{c}$, k, κ resulting from Eqs. $\hat{c} = (1 - \hat{q})^{1/k}$, (12), and (13), which fulfill the parameter constraints. Equipped with these solutions, we then evaluate, which parameter combination minimizes our MM estimator. Further details of our estimation procedure are given in the Appendix.

3.3 Estimation results

Applying the MM estimator to our dataset gives the parameter values reported in the upper panel of Table 2, with bootstrapped standard errors from 50 replications in parentheses. The estimate for ν_1 is larger than zero at a 5 percent level of significance. This indicates that the dependency of the two technology parameters z and κ in our model is empirically relevant. A closer look at the numerical solutions reveals that accounting for the dependency of the two technology parameters is indeed crucial for our estimation, since we do not find a single parameter combination with $\nu_1 = 0$ that fulfills the various constraints of our model. The value of σ is slightly lower than the one structurally estimated by Egger et al. (2013), relying on firm-level information for five European countries, but in the range of the parameter estimates of Broda and Weinstein (2006). Our estimate of k is higher than the estimates in Egger et al. (2013) and close to the shape parameters of the Pareto distribution in other studies (cf. Arkolakis, 2010; Arkolakis and Muendler, 2010).

Regarding the effective wage differential between the host and the source country of offshoring, κ, reliable estimates are not easy to find in the literature, mainly because information

Table 2: Estimation results

Parameter values					
	ν_0	ν_1	σ	k	κ
Estimates	0.22 (0.01)	0.07 (0.04)	5.19 (0.55)	8.40 (1.10)	0.56 (0.04)
Targets	m_1	m_2	m_3	m_4	
Computed	0.01 (0.00)	12.48 (0.07)	0.01 (0.00)	0.78 (0.01)	
Observed	0.02 (0.00)	12.75 (0.02)	0.10 (0.02)	0.37 (0.02)	
Difference	-0.01 (0.00)	-0.27 (0.07)	-0.10 (0.02)	0.40 (0.02)	

Note: Parameters ν_0, ν_1 and σ are determined, using the minimum-distance estimator in Eq. (18), whereas parameters k and κ are pinned down by the constraints imposed by Eqs. (12) and (13). Bootstrapped standard errors (50 replications) in parentheses.

on labor costs for a large sample of countries is unavailable. However, the US Bureau of Labor Statistics provides data on hourly labor compensation costs in manufacturing industries for several countries over a longer time horizon.[20] Using information on bilateral trade from the OECD STAN Database, this allows us to construct a sample of 32 host countries of German offshoring, for which we have information on both the value of intermediate goods imports and the hourly labor compensation costs.[21] We use this information to construct an intermediate-goods-import-weighted measure of foreign labor compensation costs for the year of 2003, which amounts to 68.33 percent of the labor compensation costs reported for Germany in this year. Since offshoring in our model is low-cost seeking, whereas the sample of host countries covers high- as well as low-cost economies, we would expect the model to predict a lower κ and thus think that an estimated value of 0.56 is of reasonable magnitude. From the lower panel of Table 2, we can furthermore conclude that the minimum distance estimator does a fairly good job in targeting the variance of marginal costs as well as the mean of log revenues of non-offshoring firms. However, we underestimate the share of offshoring firms at the first decile and overestimate it at the ninth decile.

In a second step, we apply the minimum distance estimator to a model variant, in which we impose the restriction $\nu_0 = 1$ and $\nu_1 = 0$. In this case, the probability of offshoring $\Pr_z(s > 0)$

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

[21]This country sample includes 18 of the 20 biggest suppliers of German intermediate goods imports and it covers 84.19 of bilateral intermediate goods imports to Germany reported by the OECD STAN Database.

equals one if a firm makes the offshoring investments f, and hence the model enforces an outcome without overlap. In order to make the more restrictive model consistent with the observed share of offshoring firms, we lift the constraint on $\hat{c}$ imposed by the data and estimate the two model parameters $\hat{c}$ and σ. Eqs. (12) and (13) can then again be used to solve for theory-consistent values of k and κ, whereas conditions $\sigma > 1$ and $k > 2(\sigma - 1)$ confine the possible parameter space. Furthermore, to make the parameter estimates directly comparable to those reported in Table 2, we consider the same moment conditions as in the model variant with overlap.[22] The estimation results are reported in Table 3.

Table 3: Estimation results for the no-overlap case

Parameter values				
	$\hat{c}$	σ	k	κ
Estimates	0.86 (0.02)	5.91 (0.53)	9.84 (1.06)	0.86 (0.03)
Targets	m_1	m_2	m_3	m_4
Computed	0.00 (0.00)	12.71 (0.05)	0.00 (0.00)	1.00 (0.00)
Observed	0.02 (0.00)	12.75 (0.02)	0.10 (0.02)	0.37 (0.02)
Difference	-0.02 (0.00)	-0.04 (0.04)	-0.10 (0.02)	0.63 (0.02)

Note: Parameters $\hat{c}$ and σ are determined, using the minimum-distance estimator in Eq. (18), whereas parameters k and κ are pinned down by the constraints imposed by Eqs. (12) and (13). Bootstrapped standard errors (50 replications) in parentheses.

Except for $\hat{c}$, all the parameter estimates are larger than their counterparts in Table 2. In particular, the estimate for κ seems to be unrealistically high in the context of offshoring that is low-cost seeking. Compared to the model variant with overlap, the minimum distance estimator is somewhat less successful in targeting moments 1 and 4, whereas it is more successful in capturing the mean of log revenues of non-offshoring firms.

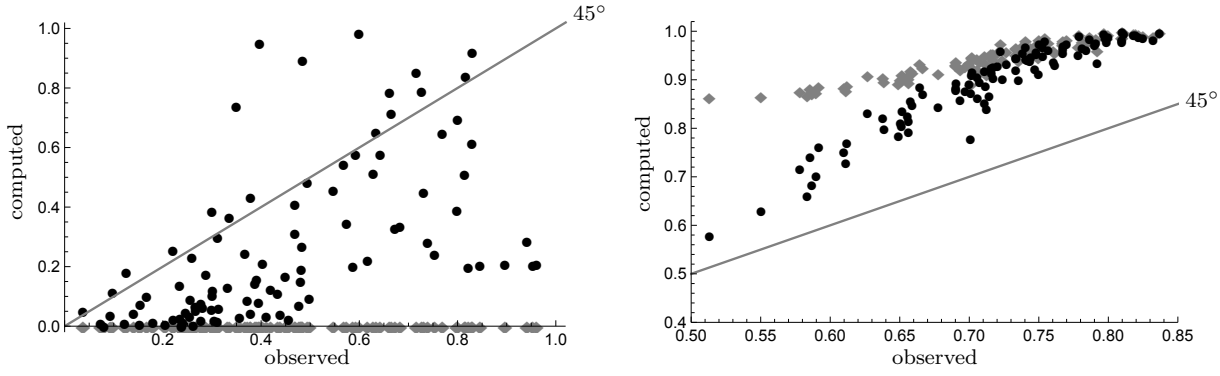
3.4 Model fit

To shed further light on the aptitude of the two model variants to capture important features of the data, we contrast computed and observed values of overlap, marginal costs, and revenues at the percentile level. Thereby, we rank firms according to their positions in the revenue distribution and construct the relevant information by averaging the firm data over an interval of

[22] Dropping the two moment conditions capturing the shares of offshoring firms does not affect our estimates.

ten millentile observations around the respective percentile positions. Following this procedure, we can determine the share of offshoring firms for each percentile in our dataset, $\widehat{\chi}_q$, and then construct the observed overlap for the various percentiles, according to $1 - |1 - 2\widehat{\chi}_q|$. This overlap measure is hump-shaped. It takes a value of 0 if either none or all firms offshore and reaches a maximum value of 1 if the number of offshoring and non-offshoring firms is equal, i.e. if $\widehat{\chi}_q = 0.5$. The theory counterpart to this overlap measure can be constructed by computing the share of offshoring firms for a certain quantile, $\chi(c_q)$, following the steps outlined in Section 3.2. The resulting value can then be used to compute a theory measure of overlap, according to $1 - |1 - 2\chi(c_q)|$.

The fit between observed and computed overlap is reported in Panel A of Figure 4. To distinguish the two model variants, we use black dots for the model with overlap and gray diamonds for the model without overlap and find that our model underestimates the overlap in the data even if we allow for arbitrary values of ν_0, ν_1 that accord with $\nu_0, \nu_1 \geq 0$ and $\nu_0 + \nu_1 \leq 1$. On average, the overlap in our preferred model amounts to 0.26, whereas the overlap in the data equals 0.44, when relying on percentile information. If we set $\nu_0 = 1$ and $\nu_1 = 0$, the overlap is by construction zero and hence the downward bias more severe.



Panel A: Offshoring overlap

Panel B: Marginal costs

● Model variant with overlap ♦ Model variant without overlap

Figure 4: Model fit: Overlap and marginal costs

In Panel B of Figure 4, we contrast observed and computed marginal costs of non-offshoring firms and find that both model variants systematically overestimate the marginal costs for all percentiles, with the upward bias being more pronounced in the model variant without overlap. Despite their problems in capturing the level of marginal costs, both models describe quite well the profile of marginal costs over the various revenue categories. In the Appendix, we show

that we can reconcile the observed level of marginal costs with our model, when allowing for a more flexible cost specification or when adding the mean of marginal costs as additional moment condition in minimization problem (18).

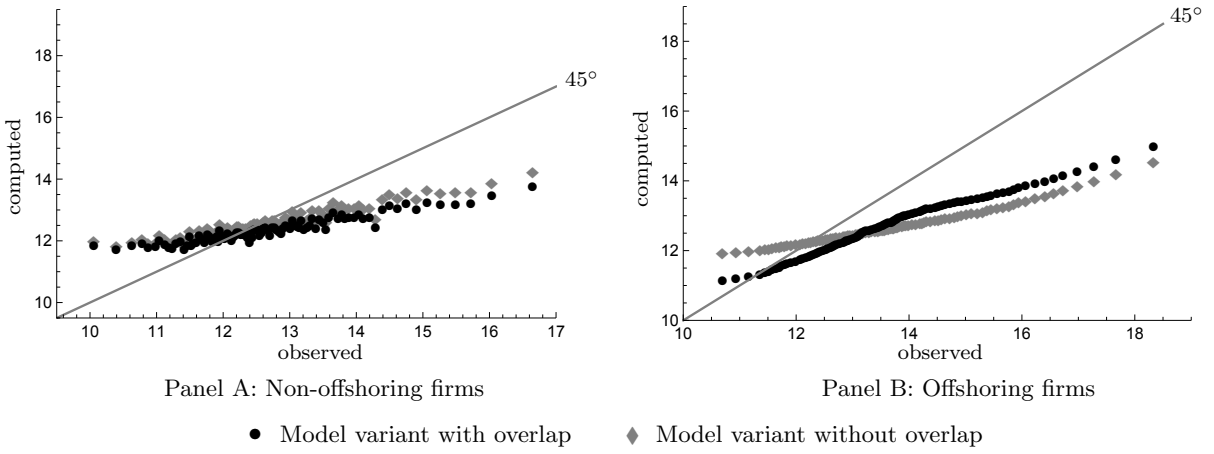


Figure 5: Model fit: Log normalized revenues

In Figure 5, we contrast observed and computed values of log revenues of non-offshoring firms (Panel A) and offshoring firms (Panel B). Computed values of revenues for non-offshorers are constructed, using the estimate of σ and the log of observed marginal costs (averaged over neighboring millentiles) as inputs in Eq. (4) and considering the firm with $\ln c_1 = -0.04$ and $\ln r_1 = 11.15$ as an anchor. From Panel A, we see that on average the two model variants do a fairly good job in capturing the observed link between marginal costs and revenues in the data. This indicates that the explanatory power of the estimated σ is fairly good. However, the two models somewhat underestimate the pronounced increase of revenues over size categories. We conduct a similar exercise for offshoring firms. This exercise is hampered, because we do not observe the marginal costs of offshorers and thus have to estimate them, using the parameter values from Tables 2 and 3, respectively, in Eq. (6) in order to learn about the marginal costs of offshorers from their position in the revenue distribution. Based on these marginal cost estimates, we then compute theory-consistent revenues of offshoring firms from Eq. (4), as outlined above. Overall, we see from Panel B that both model variants underestimate the level as well as the relatively steep increase of log revenues over size categories of offshoring firms. Since revenue information from offshoring firms has not been used for estimating the model parameters, we can interpret the results in Panel B as out of sample predictions of our model.

Figure 5 may give the impression that the model without overlap is somewhat more successful in predicting the distribution of revenues in our dataset. To see whether such concerns are warranted, we go one step further and compute the Kullback-Leibler divergence (KLD, in short) to measure the fit of our model with the observed distribution of revenues. The KLD is a relative entropy index and was introduced by Kullback and Leibler (1951) as "a measure of "distance" or "divergence" between statistical populations" (p. 79). It can be used to compute the information loss from approximating the true probability distribution of a variable ξ, $\hat{Q}(\xi)$, by probability distribution $Q(\xi)$. Thereby, the divergence between the two probability distributions captures the expected logarithmic difference between $\hat{Q}$ and Q and for discrete distributions it is given by:

$$\mathcal{D}(\hat{Q}||Q) = \sum_{\xi} \hat{Q}(\xi) \left[\ln \left(\frac{\hat{Q}(\xi)}{Q(\xi)} \right) \right], \tag{21}$$

provided that $\hat{Q}(\xi) = 0$ whenever $Q(x) = 0$ (see Gray, 1990, chapter 2). $\mathcal{D}(\hat{Q}||Q)$ is zero if the two distributions are identical and larger than zero if they are not. A smaller value of $\mathcal{D}$ indicates less information loss and thus a better approximation of the true distribution $\hat{Q}$.[23]

In our application, we associate ξ with realizations of revenues and look at size groups in order to avoid the artifact that in a small sample each revenue observation exists exactly one time. Size groups are constructed by symmetric intervals of ten millentiles around each percentile observation. This gives 99 size groups, $\xi \in \{1, ..., 99\}$, with lower and upper revenue bounds of these groups denoted by $\underline{r}_\xi$ and $\overline{r}_\xi$, respectively.[24] The true probability that a random draw from one of the size groups gives an observation in size group ξ is then $\hat{Q}(\xi) = Pr(\underline{r}_\xi \leq r \leq \overline{r}_\xi) \approx 0.01$ and the same for all ξ. To determine the corresponding value of $Q(\xi)$ from our model, we use the smallest non-offshoring firm with revenue $r_0 = 3,933.95$ and marginal cost $c_0 = 0.89$ as reference in Eq. (4) to compute theory-consistent upper and lower bounds of marginal costs that correspond to the lower and upper revenue bounds for size group ξ:

$$\overline{c}_\xi = \left(\frac{\underline{r}_\xi}{r_0} \right)^{\frac{1}{1-\sigma}} c_0 \qquad \text{and} \qquad \underline{c}_\xi = \left(\frac{\overline{r}_\xi}{r_0} \right)^{\frac{1}{1-\sigma}} c_0, \tag{22}$$

[23]The Kullback-Leibler divergence was recently used by Mrázova et al. (2016) to measure the divergence of theoretical distributions from empirical ones. These authors provide a detailed discussion about the properties of KLD and its relationship to the – for economists more familiar – (absolute) entropy of Shannon (1948).

[24]We dropped firm observations from the lowest and highest five millentiles to obtain symmetric intervals around the 99 percentiles.

respectively.[25] We can then determine the probability that a random draw of firms gives an observation from size group ξ, using the pdf from Eq. (6) together with the marginal cost bounds from Eq. (22) in $\tilde{Q}(\xi) = \int_{\underline{c}_\xi}^{\overline{c}_\xi} g_c(c)dc$. $Q(\xi)$ is then defined as the probability to draw revenues from size group ξ, conditional on drawing from one of the 99 size groups: $Q(\xi) \equiv \tilde{Q}(\xi)/\big(\sum_\xi \tilde{Q}(\xi)\big)$. Substituting the observed probability distribution $\hat{Q}(\xi)$ and the computed probability distribution $Q(\xi)$ into Eq. (21), we can finally compute

$$\mathcal{D}_w(\hat{Q}||Q) = 3.53 \ (0.18), \qquad \mathcal{D}_{w/o}(\hat{Q}||Q) = 3.79 \ (0.22) \tag{23}$$

for the model variants with and without overlap, respectively. Since the difference between $\mathcal{D}_w$ and $\mathcal{D}_{w/o}$ is small and insignificant at the 10 percent level, we must interpret the finding of $\mathcal{D}_w < \mathcal{D}_{w/o}$ with care. At least, the results in Eq. (23) do not support the conclusion that the model without overlap is more successful in capturing the distribution of revenues.

3.5 Welfare effects of offshoring

We complete the discussion in this section, by employing the parameter estimates from Tables 2 and 3 to quantify the welfare effects of offshoring in our model. For this purpose, we first note that welfare under autarky (superscript a) can be inferred from Eq. (17) by setting both $\hat{c}$ and κ equal to zero. The welfare effects of offshoring can then be computed according to $\Delta W = 100\,(W/W^a - 1)$:

$$\Delta W = 100 \left\{ \left(1 + \frac{\kappa L^*}{\tau L}\right)^{\frac{1}{\sigma-1}} \left[1 - \frac{\hat{c}^k}{\nu - \nu_1 \hat{c}} \left(\frac{\nu(\sigma-1)}{k-\sigma+1} - \frac{\nu_1 \hat{c}(\sigma-2)}{k-\sigma+2}\right) \frac{f}{f_e}\right]^{\frac{1}{1-\sigma}} - 1 \right\}. \tag{24}$$

Thereby, ΔW can be interpreted as a change in GDP per capita relative to autarky. In a next step, we combine Eqs. (5), (16), and $\Gamma_2(\kappa, \hat{c}) = 0$ to solve for theory-consistent values of f, f_e, and $\tau L/L^*$ as functions of the five parameters ν_0, ν_1, σ, k, κ, and substitute the resulting expressions into Eq. (24). This gives the welfare effects of offshoring as function of the parameter estimates in Table 2.

Following this approach, we estimate a GDP per capita stimulus from the observed exposure to offshoring that amounts to 18.90 percent, with standard error 2.73, when relying on the

[25] We use the smallest firm in the dataset as anchor to make sure that the revenues bounds of all size groups are larger than this anchor.

parameter estimates of Table 2. In contrast, the welfare gain drops to 8.73 percent, with standard error 1.39, when setting $\nu_0 = 1$, $\nu_0 = 0$, and employing the parameter estimates from Table 3. Hence, the welfare estimates from offshoring are reduced by 53.81 percent (or more than 10 percentage points) when disregarding the overlap of offshoring and non-offshoring firms in the data. This sizable gap can be explained by the difference of the κ-estimates in the two model variants, which reflects a fundamental bias from ignoring the overlap of offshoring and non-offshoring firms in quantitative trade models. Since the model without overlap associates offshoring with the most productive producers, it underestimates the (marginal) cost saving from offshoring, i.e. it overestimates the true value of κ. With the gains from offshoring being directly linked to its (marginal) cost saving effect, this leads to a downward bias in the welfare estimates, when disregarding the overlap in the data.

For an interpretation of the magnitudes of our welfare estimates, they can be put in perspective to estimates reported by other studies. An interesting point of departure in this respect is the multi-country Ricardian trade model of Eaton and Kortum (2002), which has become a benchmark in the quantitative trade literature. Eaton and Kortum compute the welfare effects of a country's movement from its observed trade openness to autarky and therefore consider a comparative static experiment of similar magnitude as ours. For Germany, they report a welfare loss of only 1.3 percent from moving to a closed economy. Alvarez and Lucas (2007) use an Eaton and Kortum (2002)-type model in a calibration exercise for 60 economies. They provide a recipe on how to use their setting for computing an upper bound of the welfare gain associated with a movement from autarky to free trade. For Germany, the upper bound of welfare gain is 19.6 percent of GDP in the overly optimistic case that all obstacles to trade are eliminated. As pointed out by Caliendo and Parro (2015) welfare estimates become significantly larger in quantitative trade models when accounting for intermediate goods. Costinot and Rodriguez-Clare (2014) analyze the role of intermediates systematically and show in an illustrative example that in the case of Germany, accounting for intermediates can lead to welfare estimates that are ten times higher than estimates from models, which do not account for intermediates. Building on a model, in which production features the assembly of tasks, we therefore think that the welfare estimates from the model variant with overlap are of reasonable magnitude.

4 Offshoring at the turn of the millennium

To understand how German offshoring has evolved over the last 25 years, we use information from the OECD STAN database and the WITS database of the Worldbank and construct a comprehensive measure of offshoring, which accounts for the import of both goods and service inputs (see the Appendix for details). Dividing the resulting measure by GDP gives the German openness to offshoring, e_{off}, which has increased from 18.03 percent in 1990 to 30.26 percent in 2014. We can now use our model to decompose this increase into changes at the intensive margin – capturing changes in the offshoring activity of incumbent offshoring firms – and the extensive margin – capturing changes in the mass of offshoring firms. To do so, we specify a theory-consistent measure of offshoring openness $e_{\text{off}} = \kappa/(\tau L/L^*)$ and compute for each year values of the exogenous effective relative domestic labor supply $\tau L/L^*$ and the endogenous variables κ and $\hat{c}$ that are consistent with the observed $\widehat{e}_{\text{off}}$ and the two implicit general equilibrium relationships $\Gamma_1(\hat{c}, \kappa) = 0$ and $\Gamma_2(\kappa, \hat{c}) = 0$.

We conduct these computations for both the model with overlap and the model without overlap and use the thus determined parameter estimates to derive theory-consistent values of offshoring openness for a counterfactual situation, in which the mass of offshoring firms stayed constant at the 1990 level. We present a detailed discussion on how we compute these variables and an overview of our parameter estimates in the Appendix and summarize the main insights from this decomposition exercise in Figure 6. The black line in this figure depicts the observed changes of German offshoring openness, whereas the solid and dashed gray lines capture the changes of offshoring that are attributed to the intensive margin by the model variants with and without overlap, respectively.

The black line shows an overall increase in German offshoring openness since the early 1990s. However, this increase has not been monotonic. There were ups and downs over the covered time span, with three notable dips in the early 1990s, the early 2000s and, most strongly, in 2009. Aside from a slight global decline in the trade to GDP ratio at the time, the first dip in offshoring openness captures two particularities of the German reunification. Eastern German producers were less inclined to offshore, and Western German producers gained access to cheap labor in the now larger domestic economy. The second dip picks up a general decline in the trade to GDP ratio in the aftermath of the dot-com crisis – maybe reinforced by a decline in the demand for cheap foreign labor after the drastic labor market reforms in Germany at the

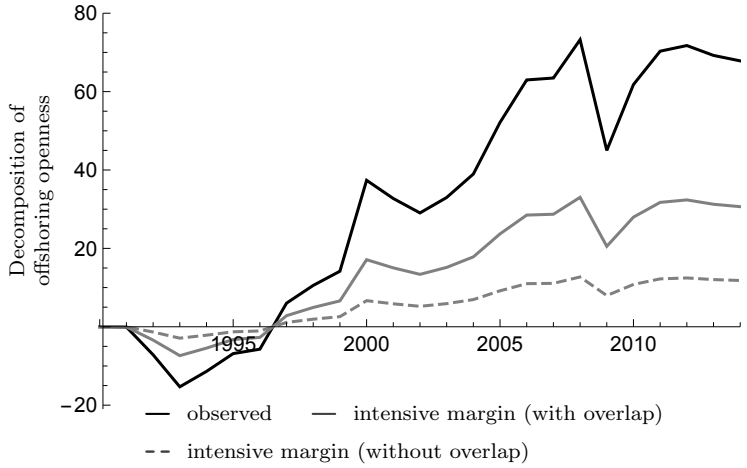


Figure 6: Changes in German offshoring openness between 1990 and 2014

beginning of the new century. Finally, the dip of offshoring openness in 2009 captures the well documented sharp decline in globalization during the financial crisis.

According to the model with overlap, both the extensive and intensive margin have played a prominent role in explaining the evolution of German offshoring openness. The intensive margin contributed 45.17 percent, with standard error 1.41, to the overall increase in German offshoring openness over the period 1990-2014. The intensive margin seems much less important, however, if one relies on the model variant without overlap, explaining only 17.41 percent, with standard error 1.94, of the increase in German offshoring openness in this case. This difference is well in line with Armenter and Koren (2015), who calibrate a quantitative trade model along the lines of Melitz (2003) with sharp sorting of firms into export mode, using US data, and compare it with an otherwise identical trade model that allows for overlap of exporters and non-exporters. In a counterfactual exercise they show that lowering the iceberg trade cost parameter leads to substantial differences of the two models regarding the relative importance of the extensive and intensive margin for explaining the increase in exporting activity, with the extensive margin being more important in the model variant without overlap.

We complete the discussion in this section by simulating the gains from offshoring over the period 1990-2014. The results of this exercise are depicted by Figure 7. In line with our insights from Section 3 the gains from offshoring are more pronounced when accounting for the observed overlap in the data. The welfare stimulus from the expansion of offshoring between 1990 and 2014 is 4.77 percent, with standard error 0.71, in the model variant with overlap and only 2.93

percent, with standard error 0.45, in the model variant without overlap. To put the size of these effects into perspective, we can contrast the offshoring gains with the overall increase in German GDP per capita between 1990 and 2014, which amounts to 38.77 percent. According to the model with (without) overlap, the increased openness to offshoring therefore explains 12.31 (7.56) percent of the overall increase in German GDP per capita since 1990.

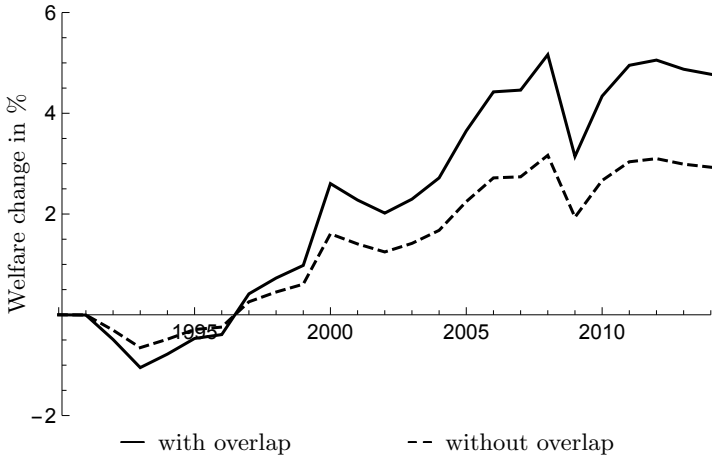


Figure 7: The welfare effects of Offshoring between 1990 and 2014

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

5 Robustness

The main insights from the analysis above are a downward bias in the welfare effects of offshoring and an exaggeration of the extensive margin of offshoring when ignoring the overlap in the data. We now analyze the robustness of this result and first consider the effect of changing the set of host countries. Since offshoring in our model is low-cost seeking, it is associated with production shifting from high-income to low-income countries. So far we have not restricted the analysis to offshoring to low-income countries, because the IAB Establishment Panel does not provide information on offshoring at the host country level. However, the dataset allows to distinguish

offshoring to EMU members from offshoring to Non-EMU countries. Hence, we can restrict the sample of host countries to the latter group and thereby increase the relative frequency of low-income economies in the dataset. We report the estimation results for the subsample of Non-EMU host countries in Panels A and B of Table 4.

There, we see that changing the composition of firms or host countries has only minor effects on most parameter estimates. However, there are notable differences to the baseline specification regarding the estimated cost saving. As expected, offshoring to low-income countries is associated with a higher cost saving, leading to a lower estimate of κ than in the baseline specification. At the same time, the share of firms that offshore to Non-EMU countries is with a value of 10.42 considerably lower than in the baseline scenario, and this causes lower welfare gains from offshoring, despite a larger cost saving in the model with overlap. The main insights from our analysis concerning the role of overlap for the size of welfare effects and the contribution of the intensive margin to observed changes in offshoring openness remain unaffected when confining the sample of host countries to Non-EMU members.

In a second extension, we address the robustness of our results regarding the chosen estimation strategy. Our baseline specification in section 3.2 has used a minimum distance approach for the estimation of the structural parameters of the model. The close relationship between our minimum distance approach and GMM is obvious from inspection of Eq. (18). The main difference is that the minimum distance approach is more restrictive.[26] To see this, we can specify a general vector function $\mathbf{\Delta}(\mathbf{o}, \mathbf{x})$ on observables $\mathbf{o}$ and structural parameters $\mathbf{x}$, which captures the data generating process. GMM requires that the moment conditions fulfill $\mathbb{E}[\mathbf{\Delta}(\mathbf{o}, \mathbf{x})] = 0$. In our MM estimation, we presume the functional form $\Delta_t(\mathbf{o}, \mathbf{x}) \equiv m_t(\mathbf{o}) - \mu_t(\mathbf{x})$, for all moment conditions $t = \{1, ..., n_m\}$, where $m_t(\mathbf{o})$ is the mean of a subset of observations $\mathbf{o}$. Hence, one can interpret the minimum-distance estimator in Eq. (18) as a GMM estimator with specific functional form of the moment conditions: $\Delta_t(\mathbf{o}, \mathbf{x}) = m_t(\mathbf{o}) - \mu_t(\mathbf{x})$ for all $t = \{1, ..., n_m\}$. Using this interpretation, one may argue to use the inverse of the variance-covariance matrix of the moment conditions, $\mathbf{V}$, for constructing weighting matrix $\mathbf{W}$, $\mathbf{W} = \mathbf{V}^{-1}$, as suggested by Hansen (1982) for GMM. Since $\mathbf{V}$ is not observable, we have to estimate it and do so by solving minimization problem (18), using the identity matrix for weighting the moment condition in a first-round estimation of parameters. We then use the thus determined parameter values as an

[26] As pointed out by Hall (2005), "the statistical framework developed by Ferguson (1958) contains many of the elements which reappeared in the GMM literature twenty-five years later" (p. 11).

Table 4: Non-EMU host countries and GMM estimation

Panel A: Non-EMU with overlap					
	ν_0	ν_1	σ	k	κ
Estimates	0.10 (0.01)	0.04 (0.03)	4.95 (0.50)	7.90 (1.01)	0.49 (0.04)
	ΔW from autarky		ΔW 1990-2014	Int. margin 1990-2014	
Effects	13.37% (1.61)		4.86% (0.56)	52.93% (2.51)	

Panel B: Non-EMU without overlap				
	$\hat{c}$	σ	k	κ
Estimates	0.79 (0.02)	5.79 (0.51)	9.60 (1.02)	0.87 (0.03)
	ΔW from autarky		ΔW 1990-2014	Int. margin 1990-2014
Effects	5.83% (0.98)		3.19% (0.50)	21.35% (2.16)

Panel C: $\mathbf{W}$ = Inverse of estimated variance-covariance matrix (with overlap)					
	ν_0	ν_1	σ	k	κ
Estimates	0.22 (0.01)	0.08 (0.04)	5.86 (0.54)	9.74 (1.08)	0.60 (0.04)
	ΔW from autarky		ΔW 1990-2014	Int. margin 1990-2014	
Effects	16.40% (2.19)		4.01% (0.60)	44.67% (1.25)	

Panel D: $\mathbf{W}$ = Inverse of estimated variance-covariance matrix (without overlap)				
	$\hat{c}$	σ	k	κ
Estimates	0.86 (0.01)	5.91 (0.44)	9.84 (0.99)	0.86 (0.02)
	ΔW from autarky		ΔW 1990-2014	Int. margin 1990-2014
Effects	8.73% (0.96)		2.93% (0.32)	17.41% (1.81)

Note: Bootstrapped standard errors (50 replications) in parentheses.

input to compute the four theory moments $\mu_t(\cdot)$ and determine the difference between the thus computed moments and the data on the observational level. The difference between observations and computed moments gives observation-specific residuals, which we use to construct the variance-covariance matrix $\widehat{\mathbf{V}}$.

Following this procedure, we estimate

$$\widehat{\mathbf{V}}_w = \begin{pmatrix} 0.001 & -0.006 & 0 & 0 \\ -0.006 & 1.427 & 0 & 0 \\ 0 & 0 & 0.093 & 0 \\ 0 & 0 & 0 & 0.233 \end{pmatrix}, \quad \widehat{\mathbf{V}}_{w/o} = \begin{pmatrix} 0.001 & -0.006 & 0 & 0 \\ -0.006 & 1.440 & 0 & 0 \\ 0 & 0 & 0.093 & 0 \\ 0 & 0 & 0 & 0.233 \end{pmatrix} \tag{25}$$

for the model variants with and without overlap, respectively. It is notable that the variance-

covariance matrices look very similar for the two model variants, with off-diagonal entries of $\widehat{\mathbf{V}}$ being zero in most cases. The reason for the latter is that the share of offshoring firms at a decile position, which is constructed from observations in the neighborhood of the decile (see above), does not covary with other observations. On the one hand, the sets of firm observations used for constructing the shares of offshoring firms at different deciles are disjoint. On the other hand, the residuals within the set of firm observations used for constructing the share of offshoring firms at a certain decile are constant by construction if one looks only at non-offshoring firms. Hence, the respective residuals at the observational level do not covary with the residuals from the variance of marginal costs or the log of revenues.

We invert matrix $\widehat{\mathbf{V}}$, check that the resulting weighting matrix $\mathbf{W}$ is positive semidefinite, and then use the thus determined weighting matrix in a second-round estimation of parameters. Panels C and D of Table 4 summarize the results of this two-step estimation approach for the model variants with and without overlap, respectively. There, we see that relying on the inverse of the variance-covariance matrix of moment conditions for specifying $\mathbf{W}$ slightly increases the estimate of κ in the model variant with overlap and therefore lowers the welfare gain attributed to the observed exposure to offshoring. The finding that the welfare effects of offshoring are considerably lower in the model variant without overlap is, however, unaffected by changes in the weighting matrix. The downward bias in the estimate for the welfare gain associated with a movement from autarky to the observed level of offshoring amounts to 46.77 percent and is therefore almost as high as in the baseline specification. Furthermore, our previous insight that ignoring the overlap leads to an exaggeration of the extensive margin of offshoring for explaining the increase of German offshoring openness over the period 1990-2014 remains valid when choosing the alternative estimation strategy.[27]

6 Concluding remarks

This paper presents a model, in which firms differ in the number of tasks they perform in the production process and the share of tasks they can offshore to a low-cost host country. Specific realizations of these two technology parameters are the outcome of a lottery and their

[27] In the Appendix, we provide further robustness checks, in which we address to what extent pooling over manufacturing and non-manufacturing firms is justified and how our results change when considering unweighted establishment data. There, we also study how the fit of our model with the observed level of marginal costs can be improved.

distributions are interdependent. More specifically, we assume that firms, which perform more tasks, have a higher probability that at least some of their tasks are offshorable. Marginal production costs decline in the number of tasks performed and the share of tasks offshored. Offshoring is subject to fixed and variables costs, and not all firms find it attractive to make the investment into offshoring. This gives a model of heterogeneoues firms, in which some but not all firms of a certain (cost or) revenue category conduct offshoring, with the share of offshoring firms increasing in revenues.

In an empirical exercise, we estimate key parameters of the model with a method of moments approach, using firm-level data from Germany. Based on the parameter estimates, we show that access to offshoring has increased welfare in Germany by 18.90 percent. This welfare estimate is more than 50 percent higher than in an otherwise identical model without overlap. The reason for this sizable gap is that a model without overlap associates offshoring with high-productivity firms and thus with firms, which by assumption require just a small marginal cost saving for finding production shifting to a low-cost country attractive. Furthermore, in a decomposition analysis we show that the increase in German offshoring over the period 1990-2014 was to a large extent driven by an increase along the intensive margin, i.e. by an expansion of offshoring by incumbent offshoring firms. This differs from the decomposition in the model without overlap, where the extensive margin, i.e. the role played by the increase in the number of offshoring firms, is exaggerated. We show that the two main insights of a downward bias in the welfare effects and the exaggeration of the extensive margin of offshoring when ignoring the overlap in the data are robust to changes in the composition of host countries and the chosen estimation strategy.

Elaborating on two important biases that materialize when ignoring the overlap of offshoring and non-offshoring firms in the data, we hope to provide a stimulus for future research on the quantitative effects of offshoring. A promising avenue for extending the analysis in this paper is to allow for firms in the host country, which in the interest of tractability have been excluded in this paper. Such an extension would shed light on the crowding out of local production by foreign labor demand of offshoring firms and would provide a framework for a rigorous welfare analysis in the host country of offshoring. An analysis along these lines would thus be informative to what extent the welfare estimates in the host country are biased when ignoring the overlap in the data and thereby complement the analysis in this study.

References

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

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

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

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

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

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

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

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

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

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

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

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

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

Broda, C. and D. E. Weinstein (2006): "Globalization and the Gains from Variety," *Quarterly Journal of Economics*, 121, 541–585.

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

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

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

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

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

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

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

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

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

Gray, R. M. (1990): *Entropy and Information Theory*, Springer, New York.

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

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

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

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

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

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

Helpman, E., M. J. Melitz, and S. R. Yeaple (2004): "Export Versus FDI with Heterogeneous Firms," *American Economic Review*, 94, 300–316.

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.

Kullback, S. and R. A. Leibler (1951): "On Information and Sufficiency," *The Annals of Mathematical Statistics*, 22, 79–86.

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

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

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

Mrázova, M., P. Neary, and M. Parenti (2016): "Sales and Markup Dispersion: Theory and Empirics," *Unpublished Manuscript*.

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

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

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

Shannon, C. E. (1948): "A Mathematical Theory of Communication," *Bell System Technical Journal*, 27, 379–423.

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. (6)

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

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

$$g_c^o\left(\tfrac{c}{w}\right) = \int_{b\in B} g_b^o(b) g_a\left(\tfrac{c}{bw}\right) \left|\tfrac{1}{b}\right| db = -\tfrac{w}{c\ln\kappa} \int_{b\in B} k b^{k-1}(\nu - \nu_1 b) db, \tag{A.1}$$

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

$$\begin{aligned} g_c^{o1}\left(\tfrac{c}{w}\right) &= -\frac{kw}{c\ln\kappa} \int_{c/w}^{c/(w\kappa)} \left[\nu b^{k-1} - \nu_1 b^k\right] db \\ &= -\frac{1}{\ln\kappa} \left\{ \nu \left(\frac{c}{w}\right)^{k-1} \left[\left(\frac{1}{\kappa}\right)^k - 1\right] - \nu_1 \frac{k(c/w)^k}{k+1} \left[\left(\frac{1}{\kappa}\right)^{k+1} - 1\right] \right\}. \end{aligned} \tag{A.2}$$

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

$$\begin{aligned} g_c^{o2}\left(\tfrac{c}{w}\right) &= -\frac{kw}{c\ln\kappa} \int_{c/w}^{\hat{b}} \left[\nu b^{k-1} - \nu_1 b^k\right] db \\ &= -\frac{1}{\ln\kappa} \left\{ \nu \left(\frac{c}{w}\right)^{k-1} \left[\left(\frac{w\hat{b}}{c}\right)^k - 1\right] - \nu_1 \frac{k(c/w)^k}{k+1} \left[\left(\frac{w\hat{b}}{c}\right)^{k+1} - 1\right] \right\}. \end{aligned} \tag{A.3}$$

Replacing $\hat{b}$ by $\hat{c}/w$ and adding up $g_c^d\left(\tfrac{c}{w}\right)$ and $g_c^o\left(\tfrac{c}{w}\right)$ for the two parameter domains gives the first and the second segment of the probability density function in Eq. (6). This completes the proof. *QED*

A.2 Derivation of Eq. (7)

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

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

Solving for the integral, gives

$$R_1 = Mr(1) \Bigg\{ (1-\nu) \frac{k}{k-\sigma+1} (\kappa\hat{c})^{k-\sigma+1} + \nu_1 \frac{k}{k-\sigma+2} (\kappa\hat{c})^{k-\sigma+2}$$

$$-\nu\frac{\kappa^{-k}-1}{\ln[\kappa]}\frac{1}{k-\sigma+1}(\kappa\hat{c})^{k-\sigma+1}+\nu_1\frac{\kappa^{-(k+1)}-1}{\ln[\kappa]}\frac{k}{k+1}\frac{1}{k-\sigma+2}(\kappa\hat{c})^{k-\sigma+2}\Bigg\}. \quad (A.5)$$

Thereby, $k > 2(\sigma-1)$ is sufficient to obtain a finite value of R_1. For the second segment, we can write

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

Solving for the integral establishes

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

Finally, for the third segment, we obtain

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

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

A.3 Properties of the offshoring indifference condition

Let us define

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

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

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

Accounting for $\lim_{\kappa\to 0}\hat{\alpha}(\kappa)=\infty$, $\hat{\alpha}(1)=0$, and $\hat{\alpha}'(\kappa)=[(\sigma-1)/\kappa](\kappa^{\sigma-1}-1)<0$, it follows that $\hat{\alpha}(\kappa)>0$ holds for all possible $\kappa<1$. Considering $\sigma>1$, we get $\alpha'(\kappa)<0$. From Eq. (8) we can thus compute

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

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

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

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

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 $(\sigma-1)(k-\sigma+2)>(\sigma-2)(k-\sigma+1)$ that the third term is positive as well. This implies $\partial\Gamma_1(\cdot)/\partial\hat{c}>0$. Putting together, we can conclude that if $\Gamma_1(\cdot)=0$ has an interior solution in $\hat{c}$ and κ, the relationship between $\hat{c}$ and κ established by $\Gamma_1(\cdot)=0$, can be determined according to

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

and is negative.

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

A.4 Derivation of Eq. (9)

The foreign wage bill is given by

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

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

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

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

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

Furthermore, we can compute

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

Putting together, we thus obtain

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

which, in view of Eq. (7), establishes Eq. (9). *QED*

A.5 Properties of the labor market constraint

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

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

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

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

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

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

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

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

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

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

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

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

Accounting for $\kappa = \tau w^*$ and thus $L + \kappa L^*/\tau = L + w^*L^*$, totally differentiating Eq. (17) with respect to τ gives

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

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

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

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

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

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

In a next step, we can note that

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

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

Differentiating Eq. (17) with respect to f gives

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

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

In the main text, we argue that the possibility of $dW/df > 0$ is the consequence of a deterioration in the source country's (double) factorial terms of trade. To support this argument, we can note that a (double) factorial terms of trade deterioration is reflected by an increase in κ and can determine the welfare effects that would result in a counterfactual scenario, in which κ does not change. To do so, we combine the offshoring indifference condition $\Gamma_1(\hat{c}, \kappa) = 0$ with Eq. (17) to rewrite source country welfare as follows

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

with

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

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

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

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

with

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

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

and

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

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

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

Accounting for

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

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

A.7 A model variant without overlap

To remove the overlap from the model we can set the probability of offshoring when making the f investment equal to one. More specifically, we can set $\nu_0 = 1$ and $\nu_1 = 0$. Whereas this modification does not affect Eqs. (1)-(3) and Eq. (4), it alters the indifference condition in Eq. (5), which now reads

$$\sigma f w = \left(\frac{\hat{c}}{w}\right)^{1-\sigma} r(w)\left[\frac{\kappa^{1-\sigma}-1}{(1-\sigma)\ln\kappa}-1\right], \tag{A.35}$$

and it changes the probability density function of c/w, which simplifies to

$$g_c\left(\tfrac{c}{w}\right) = \begin{cases} -\frac{1}{\ln\kappa}\left(\frac{c}{w}\right)^{k-1}\left[\left(\frac{1}{\kappa}\right)^k - 1\right] & \text{if} \quad \frac{c}{w} \leq \kappa\frac{\hat{c}}{w} \\ -\frac{1}{\ln\kappa}\left(\frac{c}{w}\right)^{k-1}\left[\left(\frac{\hat{c}/w}{c/w}\right)^k - 1\right] & \text{if} \quad \frac{c}{w} \in \left(\kappa\frac{\hat{c}}{w}, \frac{\hat{c}}{w}\right] \\ k\left(\frac{c}{w}\right)^{k-1} & \text{if} \quad \frac{c}{w} > \frac{\hat{c}}{w} \end{cases}. \tag{A.36}$$

Computing aggregate revenues in general equilibrium with $w = 1$, we obtain

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

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

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

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

To get a second link between $\hat{c}$ and κ we can make use of Eq. (9) and compute a modified *labor market condition*. Following the derivation steps from the main text, we obtain

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

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

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

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

and

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

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

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

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

and

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

This completes the proof. *QED*

B Empirical appendix

B.1 The definition of tasks

The 2006 BIBB/BAuA Employment Survey reports different workplace activities (Tätigkeiten) at the worker lever in addition to common occupation codes. We make use of 28 different activities which are summarized in Table B.1:

Table B.1: Task definitions

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

B.2 Details for the implementation of our estimation strategy

To construct the parameter space, we pick in a first step different combinations of ν_0, ν_1 from the unit interval in step length 0.01. Due to constraints $\nu_0, \nu_1 \geq 0$ and $\nu_0 + \nu_1 \leq 1$, this gives 5,148 possible (ν_0, ν_1)-pairs, for which we then determine the corresponding levels of k. To do so, we identify the marginal offshoring establishment, i.e. the offshoring producer with the smallest revenue in the data, and count all non-offshoring establishments with revenues lower than that. This gives $\hat{q} = 0.003$ and confirms the insight from Figure 1 that offshoring is prevalent in all size categories of German establishments. Using $\hat{c} = (1 - 0.003)^{1/k}$ in Eq. (12) and the share of offshoring firms computed from the weighted the data, $\hat{\chi} = 0.23$, we can then determine k for any combination of ν_0 and ν_1. Thereby, we restrict the possible solutions for k to values larger than 10^{-7}, because a positive value of k is enforced by our parameter constraints. Depending on the specific combination of parameters ν_0 and ν_1, the solutions for parameter k can vary between a low value of 0.31 and a high value of 103.71.

In a second step, we determine the possible values of σ in step length 0.01 that fulfill the parameter constraints $\sigma \geq 1 + 10^{-7}$ and $\sigma \leq k/2 + 1 - 10^{-7}$. For all possible combinations of the three parameter values ν_0, ν_1, σ (and the corresponding values of $\hat{c}$ and k), we then solve for $\kappa \in (0, 1)$ by equating Eq. (13) at $\widehat{R^d/R} = 0.43$. Although this problem does not have an interior solution for all possible parameter combinations, we can still fill $537,701$ of the $643,863$ cells of the parameter space. For these cells, we find a solution that fulfills all relevant parameter constraints, and we use the thus determined parameter values as input to construct the computed moments, which can then be combined with the reduced form parameter estimates to determine, which parameter combination minimizes the MM estimator.

B.3 Background material for Section 4

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

offshoring stayed constant over this period.

Table B.2 provides details on how we construct service offshoring in 2008. Starting point are the three broad (upper-tier) categories *Transportation*, *Travel*, and *Other Services*. For these categories, we determine the share of service imports associated with offshoring. In the case of transportation, these are all services not associated with the transportation of passengers. Since the disaggregated data does not add up to the value reported at the aggregate level, we introduce symmetric scaling factors to make the data reported at the disaggregated level consistent with the data reported at the aggregate level. Traveling is not related to the idea of offshoring in our model, and we therefore set its offshoring content equal to zero. Finally, we subtract life insurance and pension funding, personal cultural, and recreational services, as well as government services to obtain a measure for service offshoring in the last category. Again, we use scaling factors to make the disaggregated data consistent with the data at the aggregate. Adding up over all categories, we associate 63.69 percent of the service imports in 2008 with offshoring. We assume that this factor stays constant over time and multiply yearly information on total service imports with 0.64 to compute time-varying values of service offshoring.

Table B.2: Constructing a measure for service offshoring (BSI)

Title	Category	Service Imports	Service Offshoring	Share
Transportation	205	64,998.70	51,895.63	0.80
Travel	236	91,691.80	0.00	0.00
Other services	981	200,653.20	175,696.09	0,88
Total	200	357,343.69	227,591.72	0.64

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

The average openness observed in the data for the years 1999, 2001, and 2003, for which we have information on the offshoring status of German firms, equals $\widehat{e}_{\text{off}} = 0.23$. A theory-consistent measure of offshoring openness can be computed according to $e_{\text{off}} = \kappa/(\tau L/L^*)$, where $\tau L/L^*$ can be expressed as function of the parameter estimates in Table 2, according to Eqs. (5), (16), and $\Gamma_2(\kappa, \hat{c}) = 0$. For the model variants with and without overlap we compute an offshoring openness of 0.46 and 0.36, respectively. This implies that our model overestimates the average offshoring openness of Germany, with the upward bias being more pronounced in the model variant with overlap.

One may suspect that the downward bias in the welfare effects of offshoring when ignoring the overlap simply reflects that the model with overlap predicts larger offshoring openness. To see whether such concerns are justified, we can recompute the welfare effects from the previous section for a counterfactual situation, in which the offshoring openness in the model concurs

with the data. We can do so, by adjusting the three parameters $\tau L/L^*$, κ, and $\hat{c}$ to align computed and observed values of e_{off} subject to the offshoring indifference condition and the labor market constraint imposed by our model.[28] The welfare gains of offshoring evaluated at the thus determined parameter values amount to 14.06 and 6.50, respectively. Unsurprisingly, the welfare effects go down, when the offshoring openness is lower. However, the insight that the model variant without overlap underestimates the welfare effects of offshoring by more than 50 percent remains unaffected.

Table B.3 displays the parameter estimates (Columns 2-4), the observed changes in offshoring openness (Column 5), the intensive margin of changes in offshoring openness (Column 6) and the welfare effects of offshoring (Column 7) for the model variant with and without overlap, respectively. To determine the parameters in Columns 2-4, we proceed as described above and set $\tau L/L^*$, κ, and $\hat{c}$ to make them consistent with the observed offshoring openness and general equilibrium constraints imposed by our model. To determine the intensive margin of offshoring openness, we can first use Eq. (9) together with $L + w^*L^* = R$ to express the domestic wage bill as a function of aggregate revenues R:

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

Combining this with w^*L^* from Eq. (9), we can compute

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

for the model variant with overlap. The respective expression for the model variant without overlap can be obtained by setting $\nu_0 = 1$ and $\nu_1 = 0$ in Eq. (A.46). Evaluating the right-hand side at κ_t, $(\tau L/L^*)_t$, $t \in \{1990, ..., 2014\}$, and $\hat{c}_{1990}$ gives offshoring openness for the hypothetical case that the mass of offshoring firms stays constant at its value of 1990.[29] The relative changes of this measure are reported in Column 6. Finally, the welfare effects attributed to the observed changes in offshoring openness are computed, according to Eq. (24) and Eq. (A.44), respectively.

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

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

Table B.3: Parameter estimates, changes in offshoring openness, and welfare effects of offshoring for the period 1990-2014 (left panel = with overlap; right panel = without overlap)

Year	κ	c	$\tau L/L^*$	openness	int. margin	welfare effects
1990	0.68	0.89	3.75	0.00	0.00	0.00
1991	0.68	0.89	3.75	-0.07	-0.03	0.00
1992	0.68	0.88	4.08	-7.15	-3.40	-0.49
1993	0.69	0.87	4.55	-15.32	-7.36	-1.05
1994	0.69	0.87	4.31	-11.39	-5.45	-0.78
1995	0.68	0.88	4.07	-6.84	-3.25	-0.47
1996	0.68	0.88	4.01	-5.72	-2.72	-0.39
1997	0.67	0.89	3.50	6.03	2.83	0.42
1998	0.66	0.90	3.33	10.55	4.93	0.73
1999	0.66	0.90	3.20	14.17	6.60	0.98
2000	0.64	0.92	2.57	37.37	17.12	2.60
2001	0.64	0.92	2.68	32.73	15.03	2.28
2002	0.65	0.92	2.77	29.05	13.38	2.02
2003	0.64	0.92	2.68	32.99	15.15	2.30
2004	0.64	0.93	2.54	38.97	17.84	2.72
2005	0.62	0.94	2.28	52.14	23.70	3.65
2006	0.62	0.95	2.10	62.98	28.50	4.42
2007	0.62	0.95	2.09	63.47	28.71	4.46
2008	0.61	0.95	1.95	73.25	33.03	5.16
2009	0.63	0.93	2.41	45.00	20.52	3.14
2010	0.62	0.94	2.12	61.80	27.98	4.34
2011	0.61	0.95	1.99	70.32	31.74	4.95
2012	0.61	0.95	1.97	71.75	32.37	5.06
2013	0.61	0.95	2.00	69.23	31.26	4.87
2014	0.61	0.95	2.02	67.85	30.65	4.77
Average	0.65	0.92	2.91	31.73	14.36	2.23

Year	κ	c	$\tau L/L^*$	openness	int. margin	welfare effects
1990	0.91	0.78	5.06	0.00	0.00	0.00
1991	0.91	0.78	5.06	-0.07	-0.01	0.00
1992	0.92	0.77	5.47	-7.15	-1.34	-0.31
1993	0.92	0.76	6.03	-15.32	-2.91	-0.65
1994	0.92	0.76	5.75	-11.39	-2.15	-0.49
1995	0.92	0.77	5.45	-6.84	-1.28	-0.29
1996	0.92	0.77	5.39	-5.72	-1.07	-0.24
1997	0.91	0.79	4.75	6.03	1.11	0.26
1998	0.91	0.79	4.55	10.55	1.93	0.45
1999	0.90	0.79	4.39	14.17	2.59	0.61
2000	0.89	0.82	3.60	37.37	6.66	1.61
2001	0.89	0.81	3.74	32.73	5.86	1.41
2002	0.90	0.81	3.85	29.05	5.22	1.25
2003	0.89	0.81	3.73	32.99	5.90	1.42
2004	0.89	0.82	3.56	38.97	6.94	1.68
2005	0.89	0.83	3.23	52.14	9.18	2.25
2006	0.88	0.84	3.00	62.98	11.00	2.72
2007	0.88	0.84	2.99	63.47	11.08	2.74
2008	0.88	0.84	2.80	73.25	12.71	3.16
2009	0.89	0.82	3.40	45.00	7.97	1.94
2010	0.88	0.84	3.02	61.80	10.80	2.67
2011	0.88	0.84	2.86	70.32	12.22	3.04
2012	0.88	0.84	2.83	71.75	12.46	3.10
2013	0.88	0.84	2.88	69.23	12.04	2.99
2014	0.88	0.84	2.90	67.85	11.81	2.93
Average	0.90	0.81	4.01	31.73	5.55	1.37

B.4 Further robustness checks

In order to see whether pooling over manufacturing and non-manufacturing industries is justified, we can restrict our analysis to the 7,504 firm observations from manufacturing. The results of this analysis are reported in Table B.4, where Panels A and B refer to the model variants with and without overlap, respectively. There, we see that changing the composition of firms has only minor effects on most parameter estimates. However, there is a notable difference to the baseline specification regarding our κ estimate. Since the subsample of manufacturing industries features a larger than average share of offshoring firms of 30.66 percent, we obtain a lower κ-estimate when considering manufacturing firms only. This is intuitive, because a lower value of κ is associated with a higher cost saving from offshoring and hence makes offshoring attractive for a larger share of firms. Due to the higher cost saving, the welfare gain associated with a movement from autarky to the observed extent of offshoring increases to 44.34 percent in the model variant with overlap. As in the baseline specification the welfare effect of offshoring is more pronounced in the model with overlap than in the model without overlap. However, the difference in the estimated welfare effect of offshoring is not significant at the 10 percent level. Also the welfare effects attributed to the observed increase in offshoring openness between 1990 and 2014 by the model with overlap and the model without overlap, respectively, are not statistically different. However, we find in accordance with the baseline scenario that ignoring the overlap leads to an exaggeration of the extensive margin for explaining the observed increase in German offshoring openness over the last 25 years.

In the baseline estimation, we use the weighting schemes provided by the Research Data Centre in Nuremberg to make the IAB Establishment Panel representative for the true population of German firms. Since we are interested in economy-wide effects of offshoring, using observation weights is recommendable. However, one can also estimate the parameters using the unweighted firm information. The results of this exercise are reported in Panels C and D of Table B.4 for the model variants with and without overlap, respectively. Due to an overrepresentation of larger firms, the share of offshoring producers is higher when relying on the unweighted data. Accordingly, we estimate a lower κ and thus associate access to offshoring with larger welfare effects in both model variants. However, the main insight that ignoring the overlap in the data leads to a downward bias in the welfare effects of offshoring remains unaffected. The downward bias is somewhat lower and amounts to 38.68 percent when considering the unweighted firm data. The results in Table B.4 also confirm the insights from Section 4 that German offshoring between 1990 and 2014 has entailed considerable welfare gains and that disregarding the overlap of offshoring and non-offshoring firms in the data exaggerates the contribution of the extensive margin to the overall increase in offshoring openness over this period.

In a final extension, we look at two possible ways to improve the fit of our model with the observed level of marginal costs. One way to achieve this goal is to add the mean of marginal costs as additional target in the MM estimation. The results of this extension are summarized

Table B.4: Manufacturing firms and unweighted firm data

Panel A: Manufacturing with overlap					
	ν_0	ν_1	σ	k	κ
Estimates	0.30 (0.01)	0.05 (0.04)	4.19 (1.28)	6.38 (2.55)	0.39 (0.12)
	ΔW from autarky		ΔW 1990-2014	Int. margin 1990-2014	
Effects	44.34% (13.71)		6.56% (2.17)	44.64% (1.92)	
Panel B: Manufacturing without overlap					
	$\hat{c}$	σ	k	κ	
Estimates	0.85 (0.11)	4.63 (1.62)	7.27 (3.39)	0.65 (0.19)	
	ΔW from autarky		ΔW 1990-2014	Int. margin 1990-2014	
Effects	24.08% (23.43)		4.69% (9.28)	26.72% (4.61)	
Panel C: Unweighted firm data (with overlap)					
	ν_0	ν_1	σ	k	κ
Estimates	0.34 (0.00)	0.13 (0.04)	5.73 (0.38)	9.47 (0.76)	0.52 (0.03)
	ΔW from autarky		ΔW 1990-2014	Int. margin 1990-2014	
Effects	34.29% (3.50)		4.07% (0.39)	42.22% (0.62)	
Panel D: Unweighted firm data (without overlap)					
	$\hat{c}$	σ	k	κ	
Estimates	0.90 (0.01)	5.95 (0.31)	9.91 (0.77)	0.71 (0.02)	
	ΔW from autarky		ΔW 1990-2014	Int. margin 1990-2014	
Effects	21.02% (1.83)		3.28% (0.23)	26.10% (1.07)	

Note: Bootstrapped standard errors (50 replications) in parentheses.

in Panels A and B of Table B.5 for the model variants with and without overlap, respectively. There, we see that adding the additional moment condition lowers the estimates of σ, k, and κ – as well as $\hat{c}$ in the model variant without overlap. Due to the lower κ-estimate the welfare effects attributed to offshoring increase considerably relative to the benchmark specification and so does the contribution of the intensive margin to the observed increase of German offshoring openness over the period 1990-2014. Overall the results in the upper two panels of Table B.5 lend support to the main insight from the baseline specification that ignoring the overlap leads to a downward bias in the welfare effects of offshoring and to an exaggeration of the contribution of the extensive margin to the observed change in offshoring openness. However, the difference in the estimated welfare gains associated with a movement from autarky to the observed exposure to offshoring is not significant at the 10 percent level. Furthermore, targeting the mean of marginal costs changes the ranking of the two models regarding the size of welfare gains attributed to the observed increase in German offshoring openness over period 1990-2014. This indicates that

drawing conclusions from differences in the estimated κ-levels on the welfare effects of offshoring predicted by the two model variants can be unjustified.

Table B.5: Additional moment and alternative cost specification

Panel A: Targeting the mean of marginal costs (with overlap)					
	ν_0	ν_1	σ	k	κ
Estimates	0.21 (0.07)	0.08 (0.26)	2.79 (0.23)	3.59 (0.45)	0.26 (0.02)
	ΔW from autarky		ΔW 1990-2014	Int. margin 1990-2014	
Effects	43.42% (9.05)		15.30% (5.31)	53.56% (12.34)	

Panel B: Targeting the mean of marginal costs (without overlap)				
	$\hat{c}$	σ	k	κ
Estimates	0.46 (0.01)	1.95 (0.02)	1.91 (0.04)	0.47 (0.05)
	ΔW from autarky		ΔW 1990-2014	Int. margin 1990-2014
Effects	33.36% (3.93)		29.25% (2.17)	26.25% (2.85)

Panel C: Alternative cost specification (with overlap)					
	ν_0	ν_1	σ	k	κ
Estimates	0.22 (0.03)	0.07 (0.18)	5.19 (0.77)	8.40 (1.51)	0.56 (0.06)
	ΔW from autarky		ΔW 1990-2014	Int. margin 1990-2014	
Effects	17.57% (5.06)		9.39% (2.65)	65.56% (7.00)	

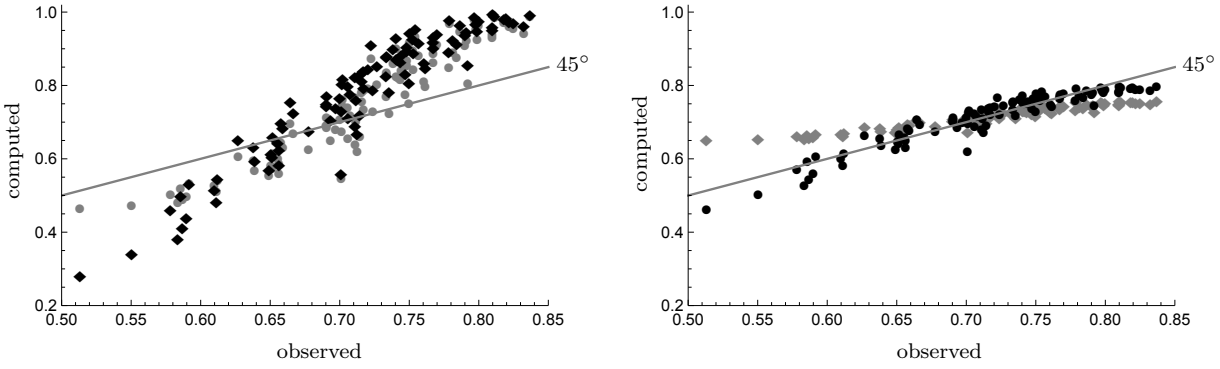
Panel D: Alternative cost specification (without overlap)				
	$\hat{c}$	σ	k	κ
Estimates	0.65 (0.01)	5.75 (0.52)	9.51 (1.20)	0.86 (0.02)
	ΔW from autarky		ΔW 1990-2014	Int. margin 1990-2014
Effects	5.59% (1.27)		9.82% (0.28)	54.41% (7.32)

Note: Bootstrapped standard errors (50 replications) in parentheses.

An alternative way to improve the fit of the model with the level of marginal costs in the data is to choose a less restrictive specification for the link between marginal costs c and task volume z, by assuming that the production of one unit of task output requires $\alpha > 0$ inputs of labor. This allows us to add α as an additional variable and the equivalence of the observed and computed mean of marginal costs as additional constraint in minimization problem (18). We put derivation details for the thus modified model to a supplement, which is available upon request, and present the estimation results in Panels C and D of Table B.5. We estimate an α-value of 0.80 and the same values as in the baseline specification for the other parameters, when looking at the model variant with overlap. In the model variant without overlap, we estimate an α-value of 0.76 and a significantly lower value of $\hat{c}$ than in the baseline specification. The main insight from our analysis that ignoring the overlap of offshoring and non-offshoring producers

in the data leads to a downward bias in the welfare effects of offshoring and an exaggeration of the extensive margin for explaining observed changes in offshoring openness remains valid if we choose a more flexible marginal cost specification. However, for the period 1990-2014 the reported differences regarding the effects on welfare and regarding the quantitative importance of the intensive margin for explaining the observed increase in offshoring openness are not statistically significant at the 10 percent level.

To complete our discussion, we finally assess to what extent targeting the mean of marginal costs or choosing an alternative cost specification has improved the fit of our model with the data. For this purpose, we can look at Figure B.1, where we see that both extensions have the intended effect of bringing the level of marginal costs predicted by the model in accordance with the observed level of marginal costs in the data. However, as illustrated in Panel A there is a trade-off between targeting the mean and the variance of marginal costs in the baseline model, so that the fit of the model with the observed changes of marginal costs over size categories deteriorates, once the mean of marginal costs is added as target in the MM estimation. Such a trade-off can be avoided when choosing a more flexible cost specification. As illustrated by Panel B, a thus extended model performs very well in capturing both the mean and variance of the marginal costs in the data.



Panel A: Targeting the mean of marginal costs

Panel B: Alternative cost specification

● Model variant with overlap ♦ Model variant without overlap

Figure B.1: Model fit: Marginal costs

S Supplement (not intended for publication)

S.1 A model variant with labour productivity $1/\alpha > 0$

In the following, we summarize the derivation details for an extended model variant with task production function $y(\omega, i) = l(i)/\alpha$, with α reflecting the labor input coefficient and thus the inverse of labor productivity in task production. In the interest of readability, we display only those equation that differ from the baseline model. In particular, Eqs. (1) and (2) are not affected by the considered modification of the task production technology. However, marginal production costs are are now given by

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

instead of Eq. (3). Accordingly, marginal costs of the least productive firm with $z = 1$ are thus α and the firm that is indifferent between making the fixed investment into f or not is characterized by

$$\sigma f w = \left(\nu - \nu_1 \frac{\hat{c}}{\alpha w}\right) \left(\frac{\hat{c}}{\alpha w}\right)^{1-\sigma} r(\alpha w) \left[\frac{\kappa^{1-\sigma} - 1}{(1-\sigma)\ln\kappa} - 1\right], \tag{S.2}$$

with $\hat{c} \equiv \alpha(1 - \hat{z})w$. Following the derivations steps of the baseline model, we can express the pdf of normalized marginal production costs as follows:

$$g_c\left(\tfrac{c}{w}\right) = \begin{cases} \left(1 - \nu + \nu_1 \frac{c/w}{\alpha}\right)\left(\frac{1}{\alpha}\right)^k k \left(\frac{c}{w}\right)^{k-1} - \frac{1}{\ln\kappa}\left\{\nu\left(\frac{1}{\alpha}\right)^k \left(\frac{c}{w}\right)^{k-1}\left[\left(\frac{1}{\kappa}\right)^k - 1\right] - \nu_1\left(\frac{1}{\alpha}\right)^{k+1} \frac{k(c/w)^k}{k+1}\left[\left(\frac{1}{\kappa}\right)^{k+1} - 1\right]\right\} & \text{if } \frac{c}{w} \le \kappa\frac{\hat{c}}{w} \\ \left(1 - \nu + \nu_1 \frac{c/w}{\alpha}\right)\left(\frac{1}{\alpha}\right)^k k \left(\frac{c}{w}\right)^{k-1} - \frac{1}{\ln\kappa}\left\{\nu\left(\frac{1}{\alpha}\right)^k \left(\frac{c}{w}\right)^{k-1}\left[\left(\frac{\hat{c}/w}{c/w}\right)^k - 1\right] - \nu_1\left(\frac{1}{\alpha}\right)^{k+1} \frac{k(c/w)^k}{k+1}\left[\left(\frac{\hat{c}/w}{c/w}\right)^{k+1} - 1\right]\right\} & \text{if } \frac{c}{w} \in \left(\kappa\frac{\hat{c}}{w}, \frac{\hat{c}}{w}\right] \\ \left(\frac{1}{\alpha}\right)^k k \left(\frac{c}{w}\right)^{k-1} & \text{if } \frac{c}{w} > \frac{\hat{c}}{w} \end{cases}. \tag{S.3}$$

To solve for the general equilibrium, we set $w = 1$ and can then relate aggregate revenues the revenues of the marginal firm with marginal cost α according to

$$R = Mr(\alpha)\left\{\left(\frac{1}{\alpha}\right)^{\sigma-1} \frac{k}{k-\sigma+1} + \hat{c}^{k-\sigma+1}\left[\nu\left(\frac{1}{\alpha}\right)^k \frac{k}{k-\sigma+1} - \nu_1\left(\frac{1}{\alpha}\right)^{k+1} \frac{k\hat{c}}{k-\sigma+2}\right]\left(\frac{\kappa^{1-\sigma} - 1}{(1-\sigma)\ln\kappa} - 1\right)\right\}. \tag{S.4}$$

Combining Eqs. (S.2) and (S.4) with the free entry condition $R = M\sigma(f_e + \hat{c}^k f)$, we obtain the new offshoring indifference condition

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

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

Linking $R = L + w^*L^*$ with constant markup pricing, we can express the foreign wage bill as follows

$$w^*L^* = R\frac{\sigma-1}{\sigma}\frac{\hat{c}^{k-\sigma+1}\left(\nu-\nu_1\frac{\hat{c}}{\alpha}\frac{k-\sigma+1}{k-\sigma+2}\right)\frac{1+\kappa^{1-\sigma}[(1-\sigma)\ln\kappa-1]}{[(1-\sigma)\ln\kappa]^2}}{1+\hat{c}^{k-\sigma+1}\left(\nu-\nu_1\frac{\hat{c}}{\alpha}\frac{k-\sigma+1}{k-\sigma+2}\right)\left(\frac{\kappa^{1-\sigma}-1}{(1-\sigma)\ln\kappa}-1\right)}. \tag{S.5}$$

In combination with $R = L + w^*L^*$ this establishes the labor market constraint:

$$\Gamma_2(\kappa,\hat{c}) \equiv \kappa\left\{\frac{\sigma}{\sigma-1}\frac{1+\hat{c}^{k-\sigma+1}\left(\nu-\nu_1\frac{\hat{c}}{\alpha}\frac{k-\sigma+1}{k-\sigma+2}\right)\left(\frac{\kappa^{1-\sigma}-1}{(1-\sigma)\ln\kappa}-1\right)}{\hat{c}^{k-\sigma+1}\left(\nu-\nu_1\frac{\hat{c}}{\alpha}\frac{k-\sigma+1}{k-\sigma+2}\right)\frac{1+\kappa^{1-\sigma}[(1-\sigma)\ln\kappa-1]}{[(1-\sigma)\ln\kappa]^2}}-1\right\}-\frac{\tau L}{L^*}=0. \tag{S.6}$$

Following the analysis in the baseline model step by step, we can furthermore express the share of offshoring firms in the total number of firms with the same marginal cost as:

$$\chi(c)=\begin{cases}1-\left[1-\frac{1}{\ln\kappa}\frac{\nu\left[\left(\frac{1}{\kappa}\right)^k-1\right]-\nu_1\frac{k}{k+1}\frac{c}{\alpha}\left[\left(\frac{1}{\kappa}\right)^{k+1}-1\right]}{[1-\nu+\nu_1\frac{c}{\alpha}]k}\right]^{-1} & \text{if} \quad c\le\kappa\hat{c}\\ 1-\left[1-\frac{1}{\ln\kappa}\frac{\nu\left[\left(\frac{\hat{c}}{c}\right)^k-1\right]-\nu_1\frac{k}{k+1}\frac{c}{\alpha}\left[\left(\frac{\hat{c}}{c}\right)^{k+1}-1\right]}{[1-\nu+\nu_1\frac{c}{\alpha}]k}\right]^{-1} & \text{if} \quad c\in(\kappa\hat{c},\hat{c}]\\ 0 & \text{if} \quad c>\hat{c}\end{cases} \tag{S.7}$$

Adding up over all c gives the overall share of offshoring firms

$$\chi=\left(\frac{\hat{c}}{\alpha}\right)^k\left[\nu-\nu_1\frac{k}{k+1}\frac{\hat{c}}{\alpha}\right] \tag{S.8}$$

Furthermore, the fraction of revenues accruing to non-offshoring producers is

$$\frac{R^d}{R}=\frac{1-\left(\frac{\hat{c}}{\alpha}\right)^{k-\sigma+1}\left(\nu-\nu_1\frac{\hat{c}}{\alpha}\frac{k-\sigma+1}{k-\sigma+2}\right)}{1+\left(\frac{\hat{c}}{\alpha}\right)^{k-\sigma+1}\left(\nu-\nu_1\frac{\hat{c}}{\alpha}\frac{k-\sigma+1}{k-\sigma+2}\right)\left(\frac{\kappa^{1-\sigma}-1}{(1-\sigma)\ln\kappa}-1\right)}, \tag{S.9}$$

which allows us to compute the economy-wide share of imported tasks

$$\rho\equiv\frac{w^*L^*}{[(\sigma-1)/\sigma]R[1-R^d/R]}=\frac{1+\kappa^{1-\sigma}[(1-\sigma)\ln\kappa-1]}{[(1-\sigma)\ln\kappa](\kappa^{1-\sigma}-1)}=\frac{\kappa^{1-\sigma}}{\kappa^{1-\sigma}-1}-\frac{1}{(1-\sigma)\ln\kappa}. \tag{S.10}$$

Combining

$$r(\alpha)=\frac{L+w^*L^*}{P^{1-\sigma}}\left(\frac{\sigma-1}{\sigma}\right)^{\sigma-1}\left(\frac{1}{\alpha}\right)^{\sigma-1} \tag{S.11}$$

with

$$r(\alpha) = \sigma f \left\{ \frac{f_e}{f} - \left(\frac{1}{\alpha}\right)^{\sigma-1} \frac{(\hat{c}/\alpha)^k}{\nu - \nu_1 \frac{\hat{c}}{\alpha}} \left[\nu \left(\frac{k}{k-\sigma+1} - \alpha^{\sigma-1} \right) \right.\right. \\ \left.\left. -\nu_1 \frac{\hat{c}}{\alpha} \left(\frac{k}{k-\sigma+2} - \alpha^{\sigma-1} \right) \right] \right\} \frac{k-\sigma+1}{k} \alpha^{\sigma-1}, \tag{S.12}$$

we can compute welfare in the source country:

$$W = \left\{ \frac{L + \kappa L^*/\tau}{\sigma f} \left[\frac{f_e}{f} - \left(\frac{1}{\alpha}\right)^{\sigma-1} \frac{(\hat{c}/\alpha)^k}{\nu - \nu_1 \frac{\hat{c}}{\alpha}} \left[\nu \left(\frac{k}{k-\sigma+1} - \alpha^{\sigma-1} \right) \right.\right.\right. \\ \left.\left.\left. -\nu_1 \frac{\hat{c}}{\alpha} \left(\frac{k}{k-\sigma+2} - \alpha^{\sigma-1} \right) \right] \right]^{-1} \frac{k}{k-\sigma+1} \left(\frac{1}{\alpha}\right)^{\sigma-1} \right\}^{\frac{1}{\sigma-1}} \frac{\sigma-1}{\sigma} \frac{1}{\alpha}. \tag{S.13}$$

We complete the formal discussion by noting that mean and variance of the marginal costs of non-offshoring producers can be expressed as

$$Mean(c) = \frac{k}{k+1} \alpha \frac{1 - \left(\frac{\hat{c}}{\alpha}\right)^{k+1} \left(\nu - \nu_1 \frac{k+1}{k+2} \frac{\hat{c}}{\alpha} \right)}{1 - \left(\frac{\hat{c}}{\alpha}\right)^{k} \left(\nu - \nu_1 \frac{k}{k+1} \frac{\hat{c}}{\alpha} \right)}, \tag{S.14}$$

$$Var(c) = \frac{k}{k+2} \alpha^2 \frac{1 - \left(\frac{\hat{c}}{\alpha}\right)^{k+2} \left(\nu - \nu_1 \frac{k+2}{k+3} \frac{\hat{c}}{\alpha} \right)}{1 - \left(\frac{\hat{c}}{\alpha}\right)^{k} \left(\nu - \nu_1 \frac{k}{k+1} \frac{\hat{c}}{\alpha} \right)} - \left[\frac{k}{k+1} \alpha \frac{1 - \left(\frac{\hat{c}}{\alpha}\right)^{k+1} \left(\nu - \nu_1 \frac{k+1}{k+2} \frac{\hat{c}}{\alpha} \right)}{1 - \left(\frac{\hat{c}}{\alpha}\right)^{k} \left(\nu - \nu_1 \frac{k}{k+1} \frac{\hat{c}}{\alpha} \right)} \right]^2 \tag{S.15}$$

respectively.