Essays on Offshoring and High-skilled Migration

Dissertation
zur Erlangung des Doktrgrades
der Wirtschafts- und Sozialwissenschaftlichen Fakultät
der Erberhard Karls Universität Tübingen

vorgelegt von

Jens Wrona
aus Eisenach

Tübingen
2014

Tag der mündlichen Prüfung:	06.05.2014
Dekan:	Professor Dr. rer. soc. Josef Schmid
1. Gutachter:	Professor Dr. rer. pol. Udo Kreickemier
2. Gutachter:	Professor Dr. rer. soc. oec. Wilhelm Kohler

Acknowledgements

First and foremost, I would like to express my deep gratitude to my first supervisor Udo Kreickemeier for his invaluable guidance and constant support in writing this thesis over the past four years. Not only he is a great co-author and a reliable source of inspiration, his thorough advice and continuous encouragement were a crucial contribution to the completion of this thesis.

I am deeply indebted to Wilhelm Kohler for his constant support and his academic advice since my time as a student in Tübingen. I would like to thank him in particular for his high personal commitment and for his willingness to become the second supervisor of my thesis.

I am very grateful to Hartmut Egger, who not only is the co-author of one of the chapters contained in this thesis, but also has provided me with invaluable advice and encouragement on other parts of this thesis. Similarly, I would like to thank my co-author, Jan Hogrefe, for his empirical expertise, which was a crucial element of our successful collaboration on another chapter contained in this thesis.

Special thanks go to my colleagues, Benjamin Jung, Miriam Kohl, Nina Neubecker and Marcel Smolka, for their support and fruitful discussions on several occasions. In addition, I would like to thank the current and former research assistants at the Chair of International Economics for their reliable support and their enthusiastic commitment.

My research greatly benefited from various presentations and seminars. In this regard, I would like to thank the participants of the Annual Congress of the European Economic Association (Glasgow, 2010; Gothenburg, 2013), Annual Meeting of the European Trade Study Group (Lausanne, 2010; Copenhagen, 2011; Leuven, 2012), Jahrestagung des Verein für Socialpolitik (Göttingen, 2012; Düsseldorf, 2013), Migration: Global Development, New Frontiers (London, 2013), Conference on Asian Dynamics in a Global Context (Copenhagen, 2010), Spring Meeting of Young Economists (Mannheim, 2012), Göttinger Workshop für Internationale Wirtschaftsbeziehungen (Göttingen, 2010,2011, 2012, 2013), Globalisation and Economic Policy (GEP) Annual Postgraduate Conference (Nottingham, 2011, 2013), RIEF Doctoral Meetings in International Trade and International Finance (Milano, 2012), Mainz Workshop on Trade and Macroeconomics (Mainz, 2013), International Workshop "Economics of Global Interactions:

New Perspectives on Trade Factor Mobility and Development" (Bari, 2012), IOS/APB Summer Academy on Central and Eastern Europe (Lake Starnberg, 2012), Ruhr Graduate School (RGS) Doctoral Conference in Economics (Dortmund, 2011; Duisburg 2012), München-Tübingen International Economics Workshop (München, 2012), Tübingen-Hohenheim Christmas Workshop (Stuttgart, 2009, 2010, 2011, 2012) and Doktorandenworkshop der Universitäten Hohenheim, Tübingen und des IAW Tübingen (Stuttgart, 2012) for valuable comments and suggestions.

Last but not least, I would like to express my warm thanks to my fiancée Julia, who constantly supported me with her abundance of patience. Further thanks go to my brother Thilo, my parents, Steffi and Jürgen, as well as to my friends for their support and encouragement during the last four years.

Tübingen, December 2013

Declaration of co-authorship

I hereby declare that this thesis incorporates material that is the result of joint research, as follows:

Chapter 3 is based on joint work with Hartmut Egger and Udo Kreickemeier. The concept for the corresponding paper Egger, Kreickemeier, and Wrona (2013a,b) was developed jointly by all three authors. Both the theoretical model and the quantitative exercise were jointly and in equal shares developed/conducted by the three authors of Egger, Kreickemeier, and Wrona (2013a,b). The two parts were mutually discussed and improved, however, such that they should be regarded as joint work. The writing of the text was shared equally.

Chapter 4 is based on joint work with Jan Hogrefe. The concept for the corresponding paper Hogrefe and Wrona (2013) was developed jointly. The theoretical model was primarily developed by the author of this thesis, while the empirical analysis was primarily conducted by Jan Hogrefe. Both parts were mutually discussed and improved, however, such that they should be regarded as joint work. The writing of the text was shared equally.

Sub-chapters 5.1 to 5.4 are based on joint work with Udo Kreickemeier. The concept for the corresponding paper Kreickemeier and Wrona (2011a,b) was developed jointly. Both the theoretical model and the empirical exercise were jointly and in equal shares developed/conducted by the two author of Kreickemeier and Wrona (2011a,b). The paper was mutually discussed and improved and hence should be regarded as joint work. The writing of the text was shared equally.

Previous publication

This thesis includes four original papers that have been previously published in form of working papers or where available as unpublished manuscripts. Chapter 3 builds on CESifo Working Paper No. 4083 (cf. Egger, Kreickemeier, and Wrona, 2013a) and on the University of Tübingen Working Paper in Economics and Finance No. 50 (cf. Egger, Kreickemeier, and Wrona, 2013b). Chapter 4 builds on the University of Tübingen Working Paper in Economics and Finance No. 64 (cf. Hogrefe and Wrona, 2013). Sub-chapters 5.1 to 5.4 build on GEP Discussion Paper 11/07 (cf. Kreickemeier and Wrona, 2011a) and on the University of Tübingen Working Paper in Economics and Finance No. 1 (cf. Kreickemeier and Wrona, 2011b). Sub-chapters 5.5 to 5.8 build on an unpublished manuscript, which is online available (cf. Wrona, 2014)

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Globalisation, which often is associated with the deepening of international trade relationships, in fact is a phenomenon with multiple facets, comprising both the international integration of goods *and* factor markets. Thereby the latter kind of market integration in its traditional sense either takes the form of cross-country capital flows or international labour migration. Within the last decades these traditional arbitrage mechanisms where complemented by a third, new paradigm, to which Grossman and Rossi-Hansberg (2008, 2012) refer as “trade in task” or “offshoring”. Unlike international labour migration, which is associated with workers moving from low- to high-wage countries, offshoring refers to the shifting of jobs from high- to low-wage locations. Thereby, as Grossman and Rossi-Hansberg (2008, p. 1987) put it: “The difference between falling costs of offshoring and falling costs of immigration is that the former create rents for domestic firms – which ultimately accrue to domestic factors in the general equilibrium – whereas the latter create rents for the immigrants.”

Both phenomena, the offshoring of production steps as well as the international migration of labour, and in particular the migration of high skilled workers, became more important in absolute *and* relative terms during the last decades. Thereby proxies for international offshoring activities can be drawn from different sources. The most often used approach follows Feenstra and Hanson (1999) and proxies offshoring through some measure of imported intermediates. Following this strategy, a comparable “index of outsourcing abroad”, developed in OECD (2007), reveals increasing levels of offshoring in many OECD countries between 1995 and 2000. Using more recent data from the World Input Output Database (WIOD) Baldwin and Lopez-Gonzalez (2013) come to a similar conclusion for the period from 1995 to 2009: Although most countries are self-sufficient in terms of intermediate inputs, "supply-chain trade" in general expands and moreover shifts towards Asia and in particular China.

For the first time Ozden, Parsons, Schiff, and Walmsley (2011) provide a complete picture of bilateral global migration for the second half of the 20th century between 1960 and 2000. The data implies that the global migrant stock increased from 92 million in 1960 to 165 million in 2000. Similarly, Hanson (2010) reports that the share of individuals residing outside their country of birth increased from 2.2% of the world population in 1980 to 3.0% in 2005. The quantitative largest part of global migration thereby occurs between developing countries, which account for half of all international migration in 2000, while the fastest growing component of global migration is migration from developing to developed countries. While bilateral migration flows, disaggregated by educational attainment, are only available for a subset of country pairs (cf. Docquier, Özden, Parson, and Artuc, 2012), there is substantial evidence that the migration of high-skilled workers became relatively more important during the same time span. With respect to the relative importance of high-skilled migration, Docquier and Marfouk (2006) report a world-wide average emigration rate of skilled workers for the year 2000 of 5.4 %, more than three times larger than the average emigration rate of all workers and almost six times larger than the emigration rate of low-skilled workers. Lowell (2007) reports a net increase of two million tertiary-educated adults who migrated between developed countries for the time span from 1975 to 2005 which is equivalent to an increase of 40%. With respect to the two-way nature of high-skilled migration, Figure 3.8 in OECD (2008a) shows that the largest destination countries for high-skilled migration, the U.S. and Canada, have substantial emigration of high-skilled individuals as well. Even more remarkably, for the United Kingdom and Germany inward and outward migration of high-skilled individuals are very similar.

In this thesis offshoring and international labour migration as the two dominant vehicles for international labour market integration are explored as two separate phenomena. This strict separation is not meant to imply that there are no interesting interactions arising from both phenomena.[1] It rather provides two separate frameworks to focus on the allocational consequences of both phenomena in isolation, which allows for a clear-cut identification of the effects that offshoring and high-skilled migration have on domestic welfare and inequality. Following this logic, the thesis is structured in two parts (see Figure 1.1), with the first part being devoted to the analysis of offshoring and the second part addressing questions related to the migration of high-skilled individuals. Both parts of the thesis comprise two content-based chapters,

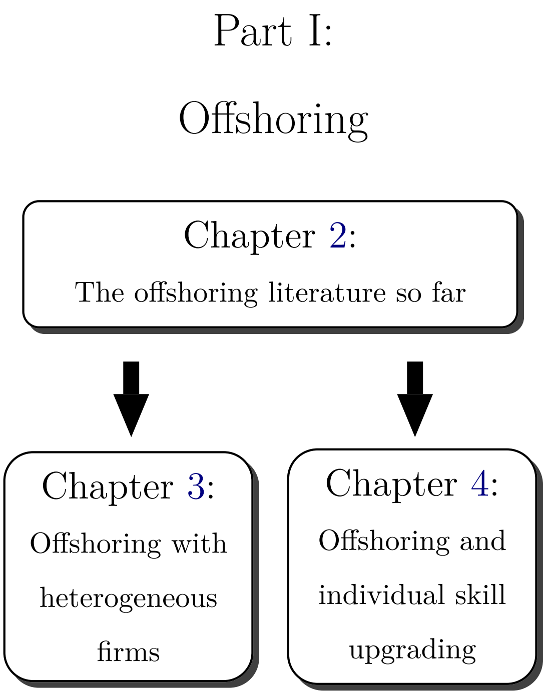


Figure 1.1: *Overview*

each addressing a specific question relating either to offshoring or high-skilled migration. The first part of the thesis in addition includes a introductory chapter providing a detailed analysis

[1] Indeed Ottaviano, Peri, and Wright (2013) recently showed that when offshore workers are specialised in the performance of tasks with an intermediate degree of complexity, which are located between the low-complexity tasks performed by immigrants and the high-complexity tasks performed by natives, immigrants are closer substitutes for offshore workers than for natives. Offshore workers may then act as a buffer between immigrants and natives absorbing potentially negative employment effects resulting from falling migration cost and increased immigration.

of offshoring within the well-known Heckscher-Ohlin framework based on the work of Kohler (2004b) and Grossman and Rossi-Hansberg (2008), as well as a short review over the theoretical offshoring literature. The necessity for such a review arises, since unlike the literature on international labour migration, the much younger offshoring literature lacks a systematic literature survey, which – to the best of my knowledge – except for Feenstra's 2008 Ohlin lecture (cf. Feenstra, 2010) so far does not exist.[2] Sticking to this logic, in the following the research questions raised in each of the four content-based chapters are introduced and discussed together with the central results from the respective chapter.

Chapter 3 – Offshoring with heterogeneous firms:

Motivated by the empirical findings that only the subset of the most productive firms engage in offshoring (cf. Moser, Urban, and Weder di Mauro, 2009; Paul and Yasar, 2009; Monarch, Park, and Sivadasan, 2013), Chapter 3 develops a simple two-country model of offshoring with monopolistic competition between heterogeneous firms at the (intermediate) goods market and occupational choice between entrepreneurship or employment as production worker at the labour market. Unlike in neoclassical trade models with atomistic firms (cf. Kohler, 2004a,b; Grossman and Rossi-Hansberg, 2008; Rodriguez-Clare, 2010), in this framework it is possible to differentiate between the effects of offshoring at the micro-level within single firms and at the macro-level within the aggregate economy. In making this distinction, the chapter's focus is on the endogenously derived extensive margin of offshoring between multinational and purely domestic firms. It is shown that the allocation of employment shares across these two types of firms crucially shapes the outcomes of the domestic economy in terms of welfare and inequality. Thereby it turns out that the effects of offshoring at the macro-level often point in exactly the opposite direction of what one would expect if guided by firm-level effects. Instrumental for this divergence

[2]For the literature on international labour migration several separate surveys exist. For general reviews of the migration literature the interested reader is referred to Borjas (1999) and Hanson (2009). The brain drain/gain literature is surveyed in more detail by Hanson (2010) and Docquier and Rapoport (2012). Detailed reviews on temporary and return migration are provided by Dustmann and Glitz (2011) and Dustmann (2001), respectively. Finally, Felbermayr, Grossmann, and Kohler (2012) provide a detailed survey on the links between international migration, international trade, and cross-country capital flows.

of micro- and macro-level results is the reallocation of workers between different firms and occupations. To give an example: Imagine a situation, where high-productivity firms reduce their domestic employment in favour of additional offshore employment abroad. At the micro-level domestic workers are displaced from their jobs, which – ceteris paribus – is bad for those workers that are directly affected (cf. Crinò, 2010; Ebenstein, Harrison, McMillan, and Phillips, 2013). But what are the implications for the aggregate economy? If the market for production labour is characterised by firm-level rent-sharing, with wages above the market-clearing level and aggregate unemployment, it is possible that the displacement of domestic workers through offshoring gives rise to an increase in economy-wide employment. How can this be the case? Offshoring firms are highly productive and pay high wages. If these high-wage jobs are offshored, and hence lost for the domestic economy, employment in the imperfectly competitive labour market for production workers becomes – ceteris paribus – less attractive relative to employment in alternative (competitive) labour markets without the risk of unemployment. To restore the indifference for (risk neutral) workers employed in different sectors of the economy, unemployment in the production sector has to decline, which then explains why the displacement of workers at the micro-level can brighten up the economy-wide employment prospects at the macro-level.

Apart from this illustrative example three general results are derived. If the firm-level employment effect in newly offshoring firms is unambiguously negative, offshoring reallocates domestic labour into less productive uses. Domestic jobs in highly productive firms vanish, and workers losing their jobs either choose to start their own firm (despite being of comparatively low productivity), they work for a domestic firm, or they find work in the service sector. This unfavourable effect on the resource allocation in the domestic economy constitutes a fundamental difference between offshoring and international goods trade, where standard models (cf. Melitz, 2003) unambiguously reallocate labour towards more productive firms; and the resulting increase in average industry productivity has been one of the important novel insights from this strand of literature (cf. Melitz and Trefler, 2012).

Despite the fact that source-country employment of newly offshoring firms may fall, their overall employment, revenues, and profits increase. It is shown that as a result decreasing offshoring costs increase the inequality of entrepreneurial incomes in a non-monotonic way:

Newly offshoring firms are at the top (bottom) of the profit distribution when the share of offshoring firms is low (high), and hence lower offshoring costs lead to more (less) inequality in entrepreneurial incomes. Moreover, inter-group inequality between entrepreneurs and workers is monotonically increasing in the share of offshoring firms. Both types of inequality are higher in any offshoring equilibrium than in autarky, and hence offshoring generates a superstar effect favouring the incomes of the best entrepreneurs, as found by Gabaix and Landier (2008).

In an extended framework with a more sophisticated model of the labour market, which allows to address the widespread concern that offshoring may have a negative effect on aggregate employment in the source country (cf. Geishecker, Riedl, and Frijters, 2012), there is rent-sharing at the firm level, leading to wage differentiation among production workers and to involuntary unemployment. The model variant with firm-level rent sharing and therefore firm-specific wage rates gives even more relevance to the domestic reallocation of workers who lost their job through offshoring. In line with empirical evidence for the US (cf. Crinò, 2010; Ebenstein, Harrison, McMillan, and Phillips, 2013), offshoring at early stages shifts employment from good manufacturing jobs (characterised by high wage premia, (cf. Krueger and Summers, 1988)) to bad (i.e. low paid) jobs. At the macro-level this generates new results regarding the effect of offshoring on aggregate unemployment, and on inequality within the group of production workers. In particular, it is shown that both the effect of offshoring on unemployment and the effect on intragroup inequality among production workers are non-monotonic in the share of offshoring firms, with unemployment and inequality being lower than in autarky when only few firms offshore, while the reverse is true when a large share of them does so. Given that production workers are identical ex ante, the model thus offers an explanation for the large variation in wage effects that offshoring has on workers within the same skill group (cf. Hummels, Jørgensen, Munch, and Xiang, 2013).

Chapter 4 – Offshoring and individual skill upgrading:

Shifting the focus from firms to workers, Chapter 4 allows individuals to react upon an offshoring shock through an adjustment of their skill acquisition strategies, which links offshoring to increased individual skill upgrading. To structure this idea, a small-open-economy model of offshoring in the spirit of Grossman and Rossi-Hansberg (2008) is set up. The model features two offshorable sets of tasks, which differ in their skill requirements. Unlike in standard trade models, where endowments are fixed, workers may react to a given offshoring shock by selecting into costly on-the-job training, thereby gaining abilities that are needed to perform skill-intensive high-wage tasks. Since the productivity effect of offshoring in Grossman and Rossi-Hansberg (2008) proportionally scales up wages for both task sets, the gap between these wages increases as well, rendering on-the-job training more attractive for untrained workers, who select into skill upgrading as long as the (offshoring induced) gap in wages exceeds the associated cost of skill upgrading.

The model's predictions are then tested using data from the German manufacturing, whereby it turns out that increased offshoring indeed has a positive and significant impact on the individual on-the-job training propensity of workers employed in German manufacturing between 2004 and 2006. This link holds for a number of specifications and is robust to the inclusion of various controls at the individual, firm, and industry level. After taking account of, among other things, technological change, business cycle effects, and firm-size differences, a one standard deviation higher offshoring growth at the industry level over the period 2004 to 2006 is related to an increase in the propensity to observe individual on-the-job training by between 3 to 7 percentage points.

Chapter 5 – Two-way migration between similar countries:

Motivated by the empirical regularity that many country pairs feature surprisingly balanced bilateral stocks of inward- and outward-migrants, Chapter 5 develops a simple framework for the analysis of permanent *and* temporary migration of high-skilled workers between similar countries. As a key feature of this modelling environment country asymmetries – underlying traditional theories of temporary and permanent migration – are absent. In the absence of such natural migration incentives high-skilled workers make use of costly two-way migration between identical countries as a signalling device (cf. Spence, 1973) to reveal their otherwise unobservable skills to potential (foreign) employers. The described migration mechanism thereby not only captures the balance of permanent bilateral migration stocks, but also explains the balance of temporary bilateral migration stocks in a dynamic (two-period) setting.

In traditional asymmetric-country frameworks welfare gains for temporary and/or permanent migrants follow as a consequence of workers' arbitrage between internationally unintegrated labour markets. Between symmetric countries such welfare effects should not arise, and indeed the derived welfare effects model contradict conventional wisdom in so far as all workers (including the migrants) tend to be worse off in an laissez-faire equilibrium with temporary and/or permanent migration than in an equilibrium without migration. Instrumental for the associated aggregate welfare loss is a negative migration externality, which leads to excessive temporary and/or permanent emigration in the presence of wasteful migration costs. As a consequence, aggregate production gains, which result from the more efficient matching of natives and migrants at the firm level, are eaten up by the costs of living abroad. Of course this does not mean that all migration, temporary or permanent, is socially harmful and, hence, must be restricted. Employing an omniscient, global social planner, it can be shown that, if the costs of living abroad are not too high, the socially optimal equilibrium may feature temporary and/or permanent migration, both – of course – at a smaller scale than in the laissez-faire equilibrium. Thereby the social planner solution in a static model can be introduced by an appropriately chosen emigration tax, while the implementation in an dynamic two-period setting requires a carefully chosen combination of emigration tax and return subsidy.

Chapter 6 – The political economy of high-skilled migration when inequality matters:

As the final part of my thesis Chapter 6 is devoted to the political economy of high-skilled migration and extends the work of Bougheas and Nelson (2012) towards an environment, in which workers display an aversion against (disadvantageous) inequality, as proposed by Fehr and Schmidt (1999). A democratic referendum with respect to the host country's (high-skilled) immigration policy will, hence, not only depend on whether the median voter benefits from immigration through a higher *absolute* (real) income, but also on the change of median voter' rank (i.e. its *relative* position) in the host country's (real) income distribution. Thereby, the median voter in the host country's labour-intensive sector – as in Bougheas and Nelson (2012) – benefits from the *indirect* terms of trade effect induced by skilled immigration, which at a global scale shifts resources into the host country's more efficient skill-intensive industry. To absorb the resulting expansion in the global production of the skill-intensive good the (relative) world market price of the skill-intensive good has to decline, and this is what leaves the median voter, who is employed in the host country's labour-intensive sector, better off, both in *absolute* and in *relative* terms. Apart from this indirect terms of trade effect skilled immigration has another more direct effect on individuals, who display an aversion against (disadvantageous) inequality, as it changes the composition of the host country's (real) income distribution. Since skilled workers from the source country immigrate into the top-ranks of the host country's income distribution, the median voter compares to a group of workers, whose skills and, hence, incomes are upward biased relative to a situation without migration. The increase in disutility from (disadvantageous) inequality aversion associated with this *direct* composition effect then – of course – must be set against the individual welfare gains from the *indirect* terms of trade effect and it is *a priori* not clear, which of both effects dominates. However, following Bougheas and Nelson (2012) in assuming a uniform distribution of workers skills, it can be shown, that the composition effect is always dominated by the terms of trade effect such that the median voter prefers an equilibrium with high-skilled migration over an equilibrium without migration, even if preferences are specified to reflect an aversion against (disadvantageous) inequality.

The remainder of this thesis is structured as follows: Chapter 2 serves as an introduction into the theoretical offshoring literature and as a point of departure for the following chapters. As outlined in Figure 1.1 what follows are the four content-based chapters discussed above, which are finally summarised in Chapter 7.

Chapter 2

The offshoring literature so far

Over the last decades the relocation of production steps towards low-wage locations abroad, commonly referred to as "offshoring", increasingly gained in importance. The academic literature on offshoring, although constantly expanding, is a rather young one – in absolute terms, with early contributions dating back to the beginnings of the nineties, and even more so in relative terms, when seen as an offspring of the much older literature on international trade in final products. As an inevitable consequence of the evolutionary process, that lead to the creation of what nowadays is regarded as a more or less unified theoretical framework for the analysis of offshoring, several terminologies along with different modelling approaches were introduced. Given that – to the best of my knowledge – so far no systematic review of the theoretical offshoring literature exists, this chapter provides an first overview over the topic and serves as starting point for the later analysis, to which the interested reader may resort whenever some of the upcoming result require a classification within the context of the recent offshoring literature.

The chapter is structured as follows: Section 2.1 defines the term "offshoring" and distiguishs the relocation of production steps towards low-wage locations abroad from related phenomena such as "outsourcing". In Section 2.2 a simple neoclassical model of offshoring based on Kohler (2004b) and Grossman and Rossi-Hansberg (2008) is developed. Within this framework the basic effects of offshoring on the allocation of factors as well as on the factor prices in the economy are analysed in a graphical way. Thereby the purpose of this section is twofold: On the one hand, the derived results serve as a reference point for the more general review of the offshoring literature in Section 2.3. On the other hand, shortcomings within the established neoclassical framework can be identified and, hence, provide a natural starting point for a more detailed view on specific aspects of offshoring in the subsequent Chapters 3 and 4.

2.1 The terminology of offshoring vs. outsourcing

A firm's (global) production structure can be classified along two dimensions, differentiating between the *organization* and the *location* of a single production process (cf. Antràs and Helpman, 2004; OECD, 2007; Feenstra, 2010). According to this classification, there exists a organizational choice to either perform all production steps within the boundary of the firm (integration) or to rely on external suppliers for the production of intermediate inputs (outsourcing). At the same time, there also is a location decision with the choice between production at the firm's headquarter location (domestic production) and production abroad (offshoring). The resulting four sub-cases from the interaction of both, the organizational and the locational dimension, are summarized in Table 2.1. Although used interchangeably in the earlier literature, the terms

	Organization of production:	
	Internal production (Integration)	External production (Outsourcing)
Location:		
Home (Domestic)	Domestic in-house production	Domestic outsourcing
Abroad (Offshoring)	Foreign in-house production	Foreign outsourcing

Table 2.1: *Offshoring vs. outsourcing*

outsourcing and *offshoring* in the above classification have distinct meanings. While outsourcing (dark gray shaded fields in Table 2.1) refers to the reliance on (domestic or foreign) external suppliers, offshoring (light gray shaded fields in Table 2.1) describes the geographical relocation of production steps across country boarders, which can be done either in terms of in-house production or by means of (foreign) outsourcing. Before this classification became necessary, due to a shift in research focus towards the organization of international production processes (see e.g. Antràs and Helpman, 2004)[1], the earlier offshoring literature referred to the relocation of production steps across country borders interchangeably as (international) "outsourcing" (cf. Feenstra and Hanson, 1996a,b, 1999; Katz and Autor, 1999), (international) "fragmentation" (cf. Jones, 2000; Kohler, 2003, 2004a,b; Egger and Kreickemeier, 2008) or "vertical specialization" (cf. Hummels, Ishii, and Yi, 2001).[2] Following the more recent literature (cf. Antràs, Garicano, and Rossi-Hansberg, 2006; Grossman and Rossi-Hansberg, 2008; Keuschnigg and Ribi, 2009; Feenstra, 2010; Rodriguez-Clare, 2010; Mitra and Ranjan, 2010), the term offshoring from now on is used to describe a situation, in which firms relocate parts of their production abroad irrespective of the organizational structure of the underlying production process.

2.2 A simple framework for the analysis of offshoring

This section introduces offshoring into a standard neoclassical 2×2 production model, abstracting for the moment from the by now common continuum-of-production-stages (cf. Kohler, 2004b), or alternatively continuum-of-tasks (cf. Grossman and Rossi-Hansberg, 2008, 2012), assumption. Instead it is assumed that there are just two task sets, labelled by $\widetilde{H}\,(H, H^*)$ and $\widetilde{L}\,(L, L^*)$, respectively, which require as sole inputs high- or low-skilled labour. Both task sets are decomposable into a non-offshorable tasks, H or L, and an offshorable task, H^* or L^*, and enter into the production of good $i = 1, 2$ according to $Y_i = F_i(\widetilde{H}, \widetilde{L})$. To simplify the analysis,

[1] For a review of this literature refer to section 3 in Helpman (2006) as well as to Antràs and Helpman (2008).

[2] Alternative, but less common labels for the same or at least very similar phenomena are "slicing up the value chain" (cf. Krugman, 1995), "trade in tasks" put forth by Grossman and Rossi-Hansberg (2008, 2012), and finally "global supply chains", recently introduced by Costinot, Vogel, and Wang (2012, 2013) to highlight the sequential production structure of their offshoring model.

a small-open-economy framework is chosen, which allows to blank out potential feedback effects that offshoring may have on the world market prices for goods *and* factors.[3] Denoting foreign wages by w_H^* and w_L^*, respectively, the offshorable tasks are performed abroad if the cost of doing so are sufficiently low, i.e. if $w_H \geq \tau_H w_H^*$ and $w_L \geq \tau_L w_L^*$, where $\tau_H, \tau_L \geq 1$ denote the usual iceberg-type offshoring cost. The unit-costs for the task sets, $\widetilde{H}$ and $\widetilde{L}$, are hence equivalent to $\omega_H(w_H, \tau_H w_H^*) \equiv \Omega_H w_H$ and $\omega_L(w_L, \tau_L w_L^*) \equiv \Omega_L w_L$, in which $\Omega_H \equiv \omega_H(\cdot)/w_H \leq 1$ and $\Omega_L \equiv \omega_L(\cdot)/w_L \leq 1$ denote the cost saving factors from offshoring tasks S^* and L^*. Analogously, the unit-cost for the final product, Y_i, may be written as $c_i(\Omega_H w_H, \Omega_L w_L) \equiv \gamma_j c_i(w_H, w_L)$ with $\gamma_i \equiv c_i(\Omega_H w_H, \Omega_L w_L)/c_i(w_H, w_L) \leq 1$ denoting the *total* cost saving factor from (partly) offshoring the inputs to the production of Y_i.

Within this deliberately simple offshoring model several scenarios can be explored, whereas offshoring is seen either as a *sector-specific* phenomenon (cf. Kohler, 2004b) or as a *task-specific* phenomenon (cf. Grossman and Rossi-Hansberg, 2008). Thereby the sector-specific view on offshoring implies that the relocation of production steps is possible in some but not in all sectors of the economy, while the task-specific perspective generally emphasises differences in offshorability between single tasks, which not necessarily depend on the sector, in which these production steps are performed. Both concepts of offshoring are analysed within the above 2×2 production model, taking into account the usual distinction between a small open economy that is either completely or incompletely specialised.

Beginning with an incompletely specialised small open economy, that produces both goods $i = 1, 2$ at given world market prices p_1 and p_2, offshoring is assumed to be sector-specific, and – without loss of generality – only possible in the labour-intensive sector two. Given the recursive structure of our simple 2×2 production model, wages in the free trade equilibrium without offshoring are solely pinned down by the zero-profit conditions $p_i = c_i(w_H, w_L)$ as illustrated in Figure 2.1. The introduction of offshoring into the above model impacts factor prices in the

[3] Ignoring these feedback effects may seem a bit awkward, given that wages in China and India – two of the world's major offshoring destinations – soared by 10 to 20 percent a year for the last decade according to The Economist (2013). To capture this offshoring driven convergence in relative wages, domestic *and* the foreign factor prices are endogeneised later on in Chapter 3. Thitherto, the simplifying assumption of constant and sufficiently low foreign wage rates applies.

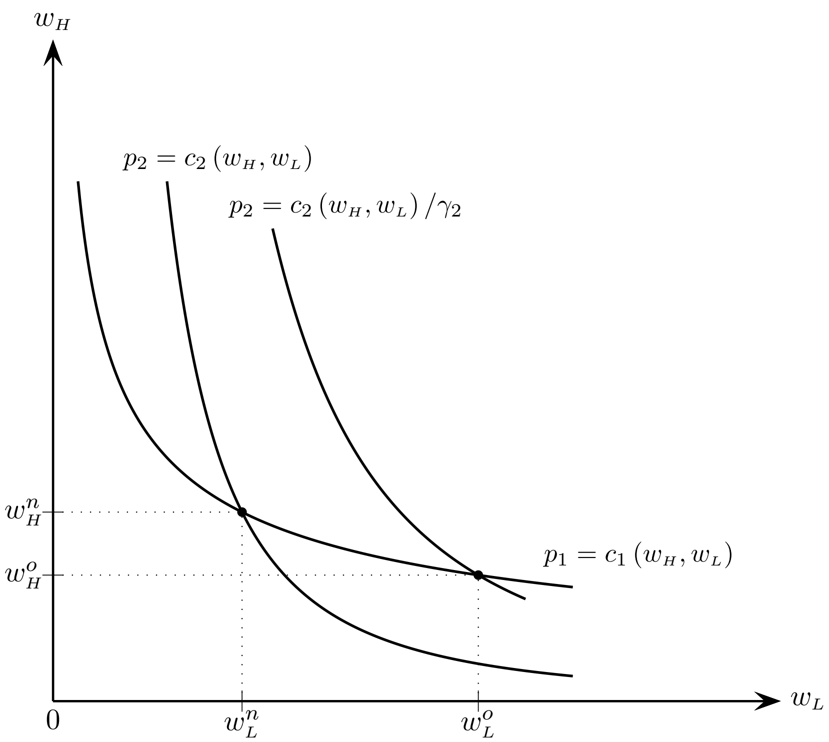


Figure 2.1: *Sector-specific offshoring in an equilibrium with incomplete specialisation*

following way: Firms in the economy's labour-intensive second sector benefit from offshoring through a reduction in their production cost by factor $\gamma_2 < 1$, which becomes possible through the access to comparably cheap foreign labour. Hence, offshoring firms appear to be more productive, and this is what Grossman and Rossi-Hansberg (2008) call the "productivity effect" of offshoring. In Figure 2.1 the productivity effect of offshoring is then reflected by an outward shift of the corresponding zero-profit condition by factor $1/\gamma_2 > 1$. This shift triggers a familiar Stolper-Samuelson mechanism (cf. Stolper and Samuelson, 1941), which causes wage gains (losses) for the factor that is more (less) intensively employed in the offshoring sector. In the chosen example (again cf. Figure 2.1) offshoring thus benefits low-skilled workers at the expense of high-skilled workers, which experience a decline in their (real) wages relative to a situation without offshoring. The fact that offshoring impacts factor prices *only* through the productivity effect is an immediate consequence of the so-called "factor price insensitivity" (cf. Leamer, 1995) in an imperfectly specialised small open economy. Workers that are displaced from their jobs in the domestic economy through offshoring thereby are absorbed by the labour market through a Rybczynski-type reallocation effect (cf. Rybczynski, 1955). To understand this reallocation effect, note, that offshoring firms necessarily set free domestic workers, which previously were employed in the now offshored tasks. Grossman and Rossi-Hansberg (2008) call this the "labour supply effect" of offshoring, since offshoring, by freeing up domestic workers, shows close resemblance to an exogenous increase in the domestic country's labour supply. The absorption of displaced workers by the domestic economy's labour market follows the Rybczynski-theorem and takes place without an adjustment in factor prices. Thereby, those workers, who disproportionately were displaced from their jobs through offshoring, are absorbed through an expansion of the sector, which disproportionately relies on this type of workers as an input into production, which renders an adjustment in (relative) factor prices superfluous.

To sum up, offshoring, when possible in only one of the economy's two sectors, benefits (hurts) the factor that is (not) intensively used in the offshoring sector through a familiar Stolper-Samuelson mechanism, while in the background displaced workers are reallocated across sectors in line with the well known Rybczynski-theorem. Thus, offshoring although beneficial for the overall economy as such, creates winners and losers along similar lines as international trade or sector-biased technological change would do. Outcomes thereby crucially depend on whether offshoring takes place in the labour- or the skill-intensive sector.

Given these insight, how does offshoring affect domestic factor prices in a scenario, in which the relocation of tasks abroad is generally possible at the same technology across both sectors? When offshoring not only happens in both sectors but is also possible at *exactly* the same technology, a surprisingly simple answer to this question exists. The impact of offshoring on unit costs of low- or high-skilled labour in such a case is the same in both sectors such that unit costs can be expressed as $\Omega_L w_L$ and $\Omega_H w_H$, respectively, with $\Omega_L, \Omega_H \leq 1$ denoting the cost-savings factor from offshoring, which is common across both sectors. In an offshoring equilibrium, the zero-profit conditions $p_i = c_i\left(\Omega_H w_H^o, \Omega_L w_L^o\right)$, as before, uniquely determine the equilibrium unit-labour cost, and from the comparison with the respective conditions $p_i = c_i\left(w_H^n, w_L^n\right)$ in the non-offshoring equilibrium, it becomes clear that *all* workers irrespective of their type benefit from offshoring as $w_L^o = w_L^n/\Omega_L$ and $w_H^o = w_H^n/\Omega_H$. Figure 2.2 illustrates this result, whereas the productivity effect of offshoring – now present in both sectors – causes an upward shift in both zero-profit conditions, which in an offshoring equilibrium support wages $w_L^o > w_L^n$ and $w_H^o > w_H^n$ in excess of what firms would pay in an equilibrium without offshoring. Thereby, displaced

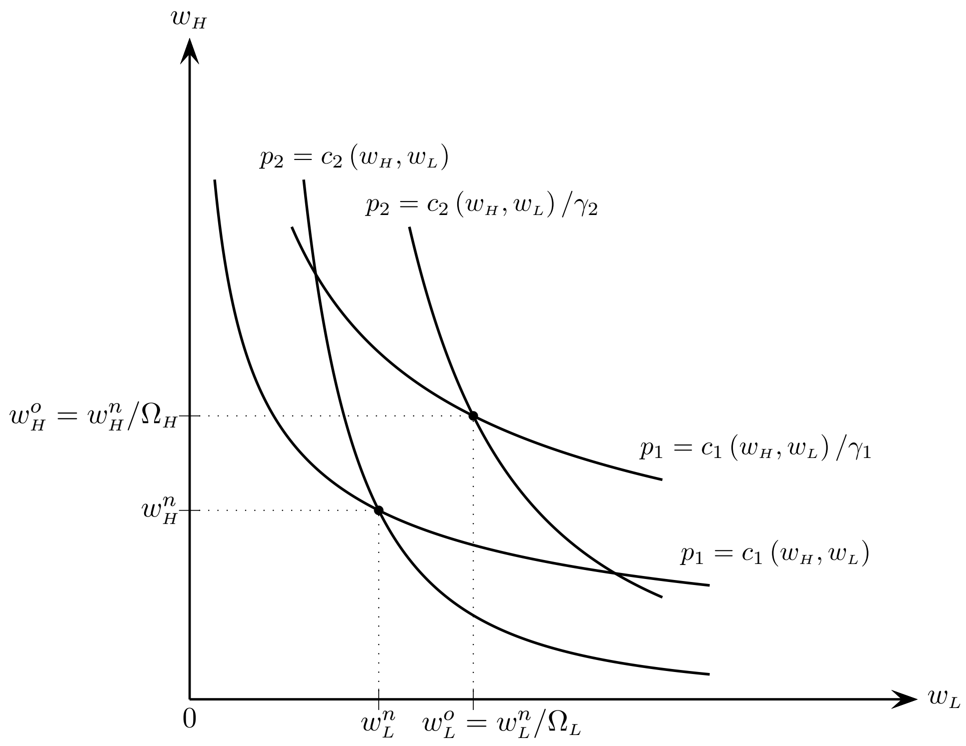


Figure 2.2: *Task-specific offshoring in an equilibrium with incomplete specialization*

workers, as it was the case before, are absorbed through a Rybczynski-type reallocation of the abundant labour towards the sector, which disproportionately relies on the respective type of workers in the production process.

Taking stock, unlike before, offshoring works to the benefit of *all* workers. This finding not only contrasts with the results from the previous scenario but also comes as a surprise, given that offshoring repeatedly has been associated with falling wages and increasing job loss fears Geishecker, Riedl, and Frijters (2012). To reconcile the different results from the previous two scenarios, it is helpful to reflect on the underlying assumptions in both cases. In the first scenario, offshoring makes firms in the offshoring sector more productive. As in the case of sector-biased technological change, the benefiting sector expands, thereby disproportionally attracting those workers that are intensively used in the sectoral production process through relatively higher wages. In the second scenario, offshoring equally lowers the cost – or, equivalently, raises the productivity – for all workers of a given type across both sectors, and hence resembles factor-biased technological change, which immediately translates into higher rewards for the benefiting factor.[4]

In both of the above scenarios, offshoring affects factor prices only through the productivity effect, while the labour supply effect *throughout* is absorbed by an inter-sectoral reallocation of workers. To rule out this smooth reallocation, let us consider now a completely specialised small, open economy, as depicted in Figure 2.3. Without the possibility of inter-sectoral reallocation of labour, (relative) factor prices respond to changes in a country's labour supply. Hence, wages under autarky depend on the domestic country's relative supply of high- to low-skilled labour $h^n \equiv \bar{H}/\bar{L}$, which in equilibrium must equal the relative labour demand (i.e. the slope $1/h$ of the unit-cost function) at wages w_L^n and w_H^n, respectively. The productivity effect of offshoring, as it was the case before, causes an upward shift in actives sector's zero-profit condition by factor $1/\gamma$. For a notionally unchanged factor intensity h^n this would imply a proportional wage increase for both low- and high-skilled workers at factors $1/\Omega_L$ and $1/\Omega_H$, respectively. However, in the absence of Rybczynski-type reallocation effects the labour supply effect of offshoring, impacts on the economy's (aggregate) skill intensity h, which makes an adjustment in (relative) factor prices unavoidable. Depending on which factor is offshored relatively more intensively, the economy's skill intensity becomes biased, which in the end benefits the factor that is offhored relatively less intensively. Whether the economy then ends up in a situation, in which *all* workers are better off, or in a scenario, which features winners *and* losers, then crucially depends on the relative size of the productivity vs. the labour supply effect of offshoring. Figure 2.3 identifies a range of $h \in [h^o_{\min}, h^o_{\min}]$, within which the economy's skill intensity supports an offshoring equilibrium that renders all workers better off compared to the autarky scenario. Thereby small deviations

[4] For a discussion of sector- vs. factor-biased technological change in neoclassical trade models see, among others, Krugman (2000) and Xu (2001).

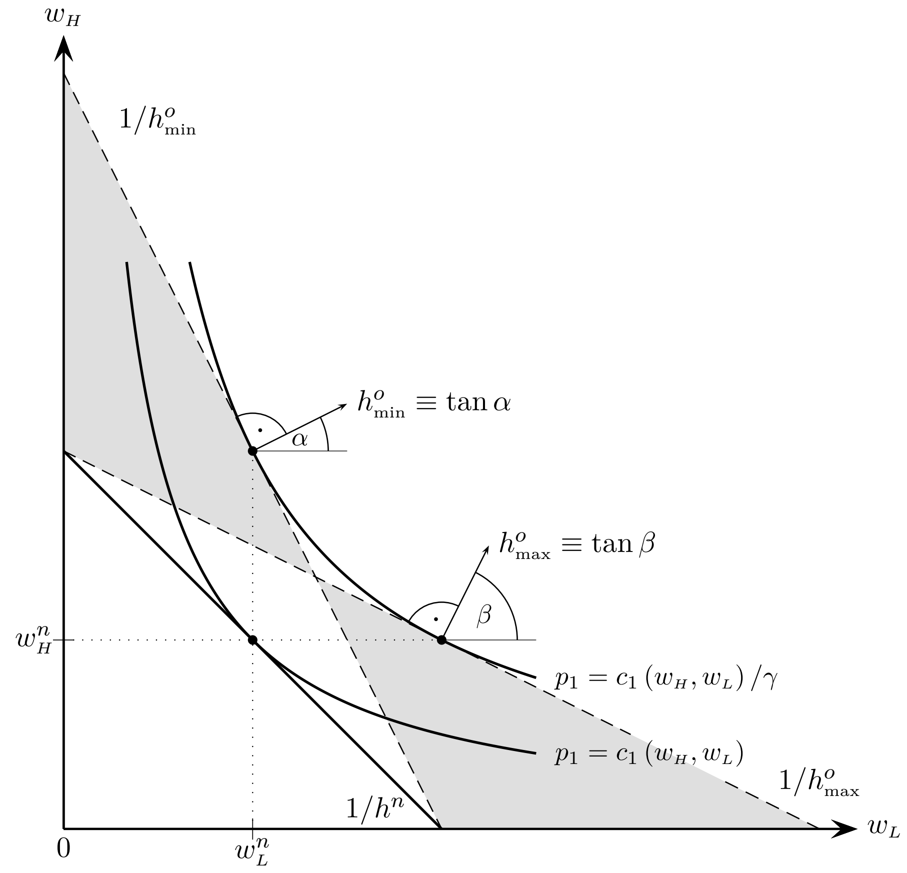


Figure 2.3: *Offshoring in an equilibrium with complete specialization*

for economy's skill intensity under autarky h^n (within the grey-shaded cone) ensure that the productivity effect of offshoring is dominant. Intuitively, this is the case when labour supply effect is relatively small or, equivalently, rather balanced across both types of factors. If instead the labour supply effect is relatively strong and/or heavily biased towards one type of factor, it becomes more likely that for a given productivity effect the economy's skill intensity under offshoring h^o deviates substantially from its respective autarky value h^n, such that the net effect of offshoring benefits (hurts) those workers that are offshored less (more) intensively.

To sum up, offshoring has two immediate effects on the domestic economy: Firms get more productive and share these gains with their employees, which in this way benefit from the productivity effect of offshoring. As the same time, workers previously employed in now offshored tasks are displaced from their jobs, which is tantamount to an increase in the domestic country's labour supply and hence summarised under the term labour supply effect. The simple model above, although highly stylised, combines both effects and offers a simple framework to analyse how offshoring impacts workers of different skills. Thereby the outcomes for the domestic economy – both in absolute *and* in relative terms – depend on two central questions: Can tasks performed by a specific factor be offshored only in a given sector or in the whole economy? And, can displaced workers be absorbed through inter-sectoral reallocation of employment shares between multiple sectors of the economy? Notwithstanding, these are not the only questions the above model may raise. Given its usefulness in illustrating the partial and aggregate effects of offshoring the highly stylised model, at the same time, features various, severe limitations in replicating the complex multidimensionality of the underlying real-world phenomenon. While some of these limitations (e.g. the heterogeneity of tasks) are addressed within the established offshoring literature (see Section 2.3), others (e.g. the heterogeneity of firms) remain (largely) unexplored and hence constitute a formidable challenge for the upcoming analysis in the sections 3 and 4.

2.3 A short review of the theoretical offshoring literature

In the theoretical literature several offshoring motives have been explored. As in the previous section offshoring often follows from a cost saving motive and exploits cross-country wage differences, which are either exogenously given (cf. Kohler, 2004a,b; Grossman and Rossi-Hansberg, 2008) or endogenised through a backward technology in the host country of offshoring (cf. Rodriguez-Clare, 2010). Alternatively, Grossman and Rossi-Hansberg (2012) motivate two-way trade in tasks between similar countries by endogenously arising wage differences, which result from the clustering of heterogeneous tasks produced under external scale economies at the national level. A similar pattern also arises in the literature on vertical specialisation (cf. Hummels, Ishii, and Yi, 2001; Yi, 2003), where trade in intermediates usually is motivated by increasing external returns to scale in the assembly of intermediate inputs (cf. Ethier, 1982). In Eckel and Egger (2009), offshoring between identical countries results as firms strategically make use of offshoring in order to undermine the bargaining position of domestic trade unions (cf. Skaksen, 2004). Finally, Grossman and Helpman (2008) explore in a model with fair wage preferences how offshoring alters the workers' fairness considerations and analyse to what extent this provides so far unexplored incentives for firms to shift production abroad.

Irrespective of the underlying offshoring motive, the effects on the offshoring firm usually are the same. As described in the previous section firms can lower their production cost by accessing cheap labour from abroad. Offshoring thus has a straightforward job relocation effect, replacing domestic through foreign employment at the firm level. The benefits from the shift of production activity to a low-cost location are ultimately reflected in a higher productivity of multinational firms. Acknowledging these firm-level effects, the theoretical offshoring literature focuses on the implications for the aggregate economy. As described in detail in the previous section a central question thereby refers to the distributional consequences that offshoring has for different kinds of workers in the domestic economy (see also Feenstra and Hanson (1996a, 1999) and Feenstra (2010) for review of this literature). Grossman and Rossi-Hansberg (2012) and Costinot, Vogel, and Wang (2012) pose a related question at a even more aggregated level and analyse how offshoring affects the global income distribution between different countries. While in Grossman and Rossi-Hansberg (2012) the production of a continuum of heterogeneous tasks can be split across two locations between which wage differentials arise as the consequence of tasks clustering in the presence of external scale economies at the national level, Costinot, Vogel, and Wang (2012) analyse multi-country supply chains, in which a country's position in the world income distribution depends on its position in the global supply chain.[5] Another concern with regard to offshoring in the open economy refers to the welfare effects of offshoring, when countries are also integrated through trade in final products. As shown by Grossman and Rossi-Hansberg (2008) and Rodriguez-Clare (2010), an expansion of domestic production capacity through the offshoring of parts from a low-wage location deteriorates the sending country's terms of trade and thus erodes the gains from trade in final goods. Finally, there also exist theoretical papers, which explore the impact of offshoring on domestic labour markets. One

[5] The structuring of offshoring processes as either "spiders", whereas all (potentially offshorable) tasks are assembled simultaneously (cf. Kohler, 2004a; Grossman and Rossi-Hansberg, 2008, 2012), versus "snakes", which require the sequential performance of production steps (cf. Costinot, Vogel, and Wang, 2012, 2013), is discussed in Harms, Lorz, and Urban (2012) and Baldwin and Venables (2013).

of them is Egger and Kreickemeier (2008), who introduce a fair-wage effort mechanism into a multi-sector traditional trade model with high-skilled and low-skilled workers to investigate the consequences of offshoring on relative wages and unemployment. Keuschnigg and Ribi (2009) study the labour market implications of offshoring in a setting with search frictions and investigate the scope of government to make offshoring Pareto improving by introducing suitable instruments of redistribution. Mitra and Ranjan (2010) consider a two-sector traditional trade model with labour market imperfection due to search frictions, and shed light on how the degree of inter-sectoral labour mobility influences the consequences of offshoring for employment and wages.

The short survey above only gives a tentative and incomplete overview, sketching out the most important cornerstones of the theoretical offshoring literature. A more detailed discussion of the literature with regard to the questions posed in the Chapters 3 and 4 is therefore given at the beginning of the respective chapters.

Chapter 3

Offshoring with heterogeneous firms

Fragmentation of production processes across country borders, leading to the offshoring of tasks that used to be performed domestically, is widely seen as a new paradigm in international trade. Public opinion in high-income countries has been very critical of this phenomenon, and much more so than of traditional forms of international trade, since it seems obvious that offshoring to low-wage countries destroys domestic jobs.[1] Academic research has drawn a picture of the effects of offshoring that invites a more nuanced view of the phenomenon than the one held by the general public. The academic literature points out that the effect of offshoring on workers in the source country is ambiguous ex ante: On the one hand, offshoring has indeed the obvious *international relocation effect* emphasised in the public discussion, as tasks that were previously performed domestically are now performed offshore, thereby harming domestic workers. On the other hand, however, there is a *productivity effect*, as the ability to source tasks from a low-wage location abroad lowers firms' marginal cost, thereby increasing overall domestic income, which benefits domestic workers, ceteris paribus.

We show in this chapter that important additional insights into the effects of offshoring can be gained by adding firm differences to the picture, thereby acknowledging the empirical

[1] As pointed out by The Economist (2009), "Americans became almost hysterical" about the job destruction due to offshoring, when Forrester Research predicted a decade ago that 3.3 million American jobs will be offshored until 2015. Using survey data from Germany, Geishecker, Riedl, and Frijters (2012) find that offshoring to low-wage countries explains about 28% of the increase in subjective job loss fears over the period from 1995 to 2007.

regularity that offshoring is highly concentrated among large firms, with many smaller firms doing no offshoring at all.[2] Both the international relocation effect and the productivity effect turn out to have new implications in the presence of firm heterogeneity, thereby jointly shaping welfare and inequality in the source country of offshoring.

To conduct our analysis, we set up a general equilibrium model that features monopolistic competition between heterogeneous firms. In many aspects, the model resembles Lucas (1978): each firm needs to be run by an entrepreneur and agents are identical in their productivity as production workers, but they differ in their entrepreneurial abilities. These abilities are instrumental for firm productivity and thus for the profit income the entrepreneur earns when becoming owner-manager of a firm. Agents are free to choose between occupations, and individual ability determines who becomes entrepreneur or production worker.[3] We extend the Lucas (1978) model to a two-country setting, and in order to introduce a stark asymmetry between the countries we assume that entrepreneurs exist in only one of them. This country ends up as the source country of offshoring, while the other country is the host country of offshoring.[4]

[2] Bernard, Jensen, Redding, and Schott (2007, 2012) show for the US that only a relatively small fraction of firms imports and that these firms systematically differ from their non-importing competitors: they are bigger, more productive, and pay higher wages. Similar evidence can be found for other countries (Wagner, 2012). This evidence is well in line with observations from a literature that looks more specifically on offshoring patterns. For instance, based on information of the IAB Establishment Panel from the Institute for Employment Research in Nuremberg, Moser, Urban, and Weder di Mauro (2009) report that only 14.9 percent of the 8,466 plants in this data-set undertake some offshoring and that, on average, offshoring firms are larger, use better technology, and pay higher wages than their non-offshoring competitors. Monarch, Park, and Sivadasan (2013) as well as Paul and Yasar (2009) report similar results for firms in the US and Turkey, respectively.

[3] Support for the occupational choice mechanism between entrepreneurship and employment as formalised in Lucas (1978) comes from matched worker-firm-owner data, which show that individuals who are unemployed (cf. Berglann, Moen, Røed, and Skogstrøm, 2011) or displaced from their job (cf. von Greiff, 2009) are more likely to select into entrepreneurship. More indirect evidence on this mechanism comes from Germany, where active labour market policies (ALMP) subsidising start-ups for unemployed (unlike other ALMP) turned out to be quite successful (cf. Caliendo and Künn, 2011).

[4] The assumption of a complete absence of entrepreneurs in the second country is not crucial for our results. Rather, it is a particularly convenient way of ensuring that the second country in the absence of offshoring would have the lower wage rate for production workers, thereby making it attractive as the destination country of offshoring. This outcome could be achieved by a less extreme assumption (e.g. by assuming that the host country

Similar to Grossman and Rossi-Hansberg (2008) and Acemoglu and Autor (2011) we model output of a firm as a composite of different tasks, and furthermore assume that only part of the tasks performed by a firm are offshorable. According to the taxonomy in Becker, Ekholm, and Muendler (2013), these are tasks that are routine (cf. Levy and Murnane, 2004) and do not require face-to-face contact (cf. Blinder, 2006). Offshoring allows to hire foreign workers for performing routine tasks at a lower wage, and this provides an incentive for firms based in the source country to shift production of these tasks abroad. This incentive is not unmitigated, since firms relocating their routine tasks abroad need to buy offshoring services, resulting in a fixed offshoring cost, and in addition shipping back to the source country the intermediate inputs produced in the host country is subject to iceberg trade costs.

As we model the production process in a similar way to Grossman and Rossi-Hansberg (2008), our model shares important features of their work. In particular, offshoring in our model and in theirs features both the international relocation effect (which Grossman and Rossi-Hansberg call "labour supply effect") and the productivity effect.[5] Since the goods market in the framework of Grossman and Rossi-Hansberg (2008) is perfectly competitive and firms are atomistic, both effects are identified in their model only in terms of their aggregate implications – the first one harming domestic workers by reducing their wage, the second one benefiting them by increasing their wage. In contrast, our framework with monopolistic competition features firms of well-defined size, and we can therefore identify the international relocation effect and the productivity effect at the firm level (with the first one leading to a reduction in domestic employment of an offshoring firm, and the second one leading to an increase), thereby allowing a direct mapping to the empirical literature using firm level data (Hummels, Jørgensen, Munch, and Xiang, 2013).

With firm heterogeneity, the firm-level effects themselves as well as their implication for the economy-wide labour allocation depend on the composition of offshoring and purely domestic firms (which itself is endogenous). If variable offshoring costs are high, only the high-productivity

has entrepreneurs, but they are less productive than in the source country), but this would add nothing interesting to our analysis, while making it considerably more complicated.

[5] Grossman and Rossi-Hansberg (2008) identify a third effect of offshoring, which materializes if the relative prices of export and import goods change in the process of offshoring. In our model with a single final good and production of this good in just one country, this terms-of-trade effect is absent.

firms benefit from shifting production of their routine tasks abroad. In this case, the firm-level productivity effect is negligible (since marginal cost savings are small due to high obstacles to international production shifting), while the international relocation effect is sizable (since all offshoring firms relocate a discrete fraction of their tasks), and therefore the firm-level employment effect in newly offshoring firms is unambiguously negative. As a consequence, offshoring unambiguously reallocates domestic labour into less productive uses. Domestic jobs in highly productive firms disappear, and workers losing their jobs in these firms either choose to start their own firm despite being of comparatively low productivity, they work for a (new or old) purely domestic firm, or they find work in the offshoring service sector. When variable offshoring costs are low, the effects are reversed: the firm-level employment effect in newly offshoring firms turns positive, and offshoring reallocates labour towards more productive firms. The potentially unfavourable effect on the resource allocation in the source country constitutes a fundamental difference between offshoring and international goods trade, where standard models with firm heterogeneity (cf. Melitz, 2003) feature an unambiguous reallocation of labour towards more productive firms; and the resulting increase in average industry productivity has been one of the important novel insights from this strand of literature (cf. Melitz and Trefler, 2012).

Despite the fact that source-country employment of newly offshoring firms may fall, their overall employment, revenues, and profits increase. We show that as a result the effect of decreasing offshoring costs on the inequality of entrepreneurial incomes is non-monotonic. The reasoning is straightforward: Newly offshoring firms are at the top of the productivity distribution when the share of offshoring firms is low, and hence lower offshoring costs in this case lead to more inequality in entrepreneurial incomes. By contrast, newly offshoring firms are at the bottom of the productivity distribution when the share of offshoring firms is high, and hence lower offshoring costs in this case lead to less inequality in entrepreneurial incomes. We also show that the effect of offshoring on inter-group inequality between entrepreneurs and workers is monotonically increasing in the share of offshoring firms. Both types of inequality are higher in any offshoring equilibrium than in autarky, and hence offshoring generates a *superstar* effect favouring the incomes of the best entrepreneurs, similar to Gersbach and Schmutzler (2007). Empirical support for this kind of superstar effect comes from Gabaix and Landier (2008), who show that small differences in managerial skills are sufficient to explain vast differences in the remuneration of US top managers, once the differences in the size of managed firms are taken into account.[6]

In the main part of this chapter, we assume that the market for production labour is perfectly competitive. While this version of our model serves the purpose well to isolate the role of firm heterogeneity in the offshoring process, we show that it is straightforward to extend the framework by using a more sophisticated model of the labour market, which allows us to address the widespread concern that offshoring may have a negative effect on aggregate employment in a country that shifts production of routine tasks to a low-wage location (cf. Geishecker, Riedl, and Frijters, 2012). In this extended version of the model, there is rent-sharing at the firm level, leading to wage differentiation among production workers and to involuntary unemployment. Interestingly, all our results from the full-employment version of the model remain qualitatively unchanged. In addition, the model variant with firm-level rent sharing and therefore firm-specific wage rates gives even more relevance to the domestic reallocation process for workers who lost their job through offshoring. In line with recent empirical evidence for the US (cf. Crinò, 2010; Ebenstein, Harrison, McMillan, and Phillips, 2013), we find that offshoring at early stages shifts employment from *good* manufacturing jobs (characterised by high wage premia, cf. Krueger and Summers, 1988) to *bad* (i.e. low paid) jobs, that for example emerge in the service sector. At the macro-level this reallocation process generates new results regarding the effect of offshoring on aggregate unemployment, and on inequality within the group of production workers. In particular, we show that both the effect of offshoring on unemployment and the effect on intra-group inequality among production workers are non-monotonic in the share of offshoring firms, with unemployment and inequality being lower than in autarky when only few firms offshore, while the reverse is true when a large share of them does so. Since all production workers are identical ex ante, our extended model offers an explanation for the large variation in wage effects

[6] In particular, Gabaix and Landier (2008) show that the sixfold increase in the remuneration of the top 500 CEOs in the US from 1980 to 2003 is well explained by the simultaneous increase in the size of firms managed by these CEOs. Although the ultimate cause for the increase in firm size is not subject of their analysis, the authors point to "greater ease of communication" (cf. Gabaix and Landier, 2008, p. 93), facilitating the global expansion of US top firms, as one possible explanation.

that offshoring has on workers within the same skill group (cf. Hummels, Jørgensen, Munch, and Xiang, 2013).

This chapter is related to the large literature that studies offshoring to low-wage countries, including the key contributions by Jones and Kierzkowski (1990), Feenstra and Hanson (1996a), Kohler (2004b), Rodriguez-Clare (2010), and, as earlier discussed in detail, Grossman and Rossi-Hansberg (2008).[7] Only few papers in the literature on offshoring consider firm heterogeneity. Antràs and Helpman (2004) were the first to analyse a firm's sourcing decision in the presence of firm heterogeneity. In their model, which features incomplete contracts, they explain the coexistence of up to four different sourcing modes (outsourcing vs. in-house production in the domestic or foreign economy, respectively) as well as the prevalence of certain sourcing patterns, when firms with different productivities self-select into these modes. Importantly, Antràs and Helpman (2004) address neither the welfare nor the distributional effects of offshoring, which are the focus of our analysis. Antràs, Garicano, and Rossi-Hansberg (2006) develop a model with team production, in which offshoring is synonymous to the formation of international teams. Individuals are heterogeneous in their skill level, and the highest-skilled individuals self-select into becoming team managers. Since individuals with higher skills are more productive in the role of a production worker as well as in the role of a manager, offshoring – by providing access to a large, relatively low-skilled foreign labour force – not only increases the incentives of workers to become managers in the source country, but also reduces the average skill level of the domestic workforce. Due to positive assortative matching between managers and workers, the top managers therefore end up being matched with workers of a lower skill level in the open economy, and hence they lose relative to less able managers. This is a key difference to the superstar effect present in our model. Davidson, Matusz, and Shevchenko (2008) consider high-

[7] In very recent work, Acemoglu, Gancia, and Zilibotti (2012) consider a Ricardian model in which offshoring induces directed technical change. With technical change favoring high-skilled workers at low levels of offshoring, this model provides a rationale for the empirical observation of rising skill premia in developed as well as developing countries. Costinot, Vogel, and Wang (2013) use a Ricardian framework with many goods and countries to study vertical specialisation of countries along the global supply chain. In Costinot, Vogel, and Wang (2012) this framework is extended to study how a country's position in the global supply chain affects the income distribution within the respective country.

skilled offshoring in a model with search frictions, in which firms can choose whether to produce with an advanced technology or a traditional technology, and workers are either high-skilled or low-skilled. Their framework is very different from ours, in that all firms hire only a single worker, and in an offshoring equilibrium they have to decide whether to do so domestically or abroad, ruling out incremental adjustments in firm level employment.[8]

The remainder of this chapter is organised as follows. In Section 3.1 we set up the model and derive some preliminary results regarding the decision of firms to offshore and its implications for firm-level profits. We also characterise the factor allocation in the open economy equilibrium and show how the share of offshoring firms is linked to the variable cost of offshoring. In Section 3.2 we analyse how changes in the offshoring costs affect factor allocation, income distribution, and welfare in our model. In Section 3.3 we present the extended version of our model that features firm-level rent sharing and involuntary unemployment. In Section 3.4 we analyse the effect of offshoring on economy-wide inequality. In Section 3.5 we use parameter estimates from existing empirical research to quantify the implications of offshoring on welfare, unemployment, and income inequality in a model-consistent way. Section 3.6 concludes our analysis with a summary of the most important results.

3.1 A model of offshoring and firm heterogeneity

We consider an economy with two sectors: A final goods industry that uses differentiated intermediates as the only inputs, and an intermediate goods industry that employs labour for performing two tasks, which differ in their offshorability. One task is non-routine and requires face-to-face communication, and it must therefore be produced at the firm's headquarters loca-

[8] There is a complementary literature that looks at offshoring between similar countries in the presence of firm heterogeneity. Amiti and Davis (2012) and Kasahara and Lapham (2013) extend Melitz (2003) and develop a model in which firms may import foreign intermediates which are then combined with domestic labour to produce the final output good. Unlike in our model, firms' sourcing decisions are driven by external increasing returns to scale in the assembly of intermediate goods (cf. Ethier, 1982) and do not follow from a cost-savings motive. In fact, the variable unit cost for imported intermediates (including variable trade cost) in these models usually exceeds the variable unit cost of domestically produced intermediates.

tion. The other task is routine and can be either produced at home or abroad. Each firm in the intermediates goods industry is run by an entrepreneur, who decides on hiring workers for both tasks. We embed the economy just described into a two-county world, where the second country differs from the first in only one respect: The second country does not have any entrepreneurs. Given our production technology, the country without entrepreneurs cannot headquarter any firms, and therefore ends up being the host country of offshoring. The other country is the source country of offshoring. Trade is balanced in equilibrium, with the source country exporting the final good in exchange for the tasks offshored to the host country. In the remainder of this section, we discuss in detail the main building blocks of our model and derive some preliminary results.

3.1.1 The final goods industry

Final output is assumed to be a CES-aggregate of differentiated intermediate goods $q(v)$:

$$Y = \left[M^{(1-\varepsilon)(\rho-1)} \int_{v \in V} q(v)^{\rho} dv \right]^{1/\rho}, \tag{3.1}$$

where V is the set of available intermediate goods with Lebesgue measure M, and $\rho \in (0,1)$ is a preference parameter that is directly linked to the elasticity of substitution between the different varieties in the production of Y: $\sigma \equiv (1-\rho)^{-1} > 1$. Parameter $\varepsilon \in [0,1]$ determines the extent to which the production process is subject to external increasing returns to scale, analogous to Ethier (1982). As limiting cases we obtain for $\varepsilon = 0$ the production technology without external increasing economies of scale, as in Egger and Kreickemeier (2012), and for $\varepsilon = 1$ the textbook CES production function with external increasing returns to scale, as in Matusz (1996). We choose Y as the *numéraire* and set its price equal to one. Profit maximisation in the final goods industry determines demand for each variety v of the intermediate good:

$$q(v) = \frac{Y}{M^{1-\varepsilon}} p(v)^{-\sigma}. \tag{3.2}$$

As will become clear in the following, the size of ε, and hence the extent of external increasing returns to scale, does not affect our results apart from those on welfare.

3.1.2 The intermediate goods industry

In the intermediate goods sector, there is a mass M of firms that sell differentiated products $q(v)$ under monopolistic competition. Each firm is run by a single entrepreneur who acts as owner-manager and combines a non-routine task, which must be performed at the firm's headquarters location in the source country, and a routine task, which can either be produced at home or abroad. We denote the non-routine task by superscript n and the routine task by superscript r. In analogy to Antràs and Helpman (2004) and Acemoglu and Autor (2011), we assume that the two tasks are inputs in a Cobb-Douglas production function for intermediate goods. Assuming that one unit of labour is needed for one unit of each task, the production function for intermediates can be written as

$$q(v) = \varphi(v) \left[\frac{l^n(v)}{\eta}\right]^{\eta} \left[\frac{l^r(v)}{1-\eta}\right]^{1-\eta}, \tag{3.3}$$

where $\varphi(v)$ denotes firm-specific productivity, $l^n(v)$ and $l^r(v)$ are the labour inputs in firm v for the production of the respective tasks, and $\eta \in (0,1)$ measures the relative weight (cost share) of the non-routine task in the production of the intermediate good.[9] Firms select into one of two categories: either they become a purely domestic firm, denoted by superscript d, or they become an offshoring firm, denoted by superscript o. The two types of firms differ with respect to the unit production cost for the routine task: For a purely domestic firm, performing the routine task onshore, this cost is simply equal to the domestic wage rate w. For an offshoring firm, hiring labour for this task in the host country, the cost is equal to the effective host country

[9] Our production function can easily be extended to account for a continuum of tasks that differ in their offshorability as in Grossman and Rossi-Hansberg (2008). Firms would then not only choose their offshoring status, but also decide on the range of tasks they relocate abroad. In Appendix A.1 we show that all offshoring firms would choose to offshore the same range of tasks, irrespective of their own productivity $\varphi(v)$. As the only additional effect in this more sophisticated model variant, a change in the cost of offshoring would not only be associated with a change in the share of firms entering offshoring, but also with a change in the range of tasks offshored by infra-marginal firms. Since the general equilibrium implications of the latter effect are well understood from Grossman and Rossi-Hansberg (2008), we focus here on the extensive margin of offshoring between rather than on the intensive margin within firms. For an extension of the Grossman and Rossi-Hansberg (2008) framework to a production technology that allows for arbitrary degrees of substitution in the assembly of tasks, see Groizard, Ranjan, and Rodriguez-Lopez (2013).

wage rate τw^*, where $\tau > 1$ represents the iceberg transport costs an offshoring firm has to incur when importing the output of the routine task from the offshore location.[10] The constant marginal costs of producing output $q(v)$ for the two types of firms are therefore given by

$$c^d(v) = \frac{w}{\varphi(v)} \qquad \text{and} \qquad c^o(v) = \frac{w}{\varphi(v)\kappa}, \qquad \text{where} \qquad \kappa \equiv \left(\frac{w}{\tau w^*}\right)^{1-\eta} \tag{3.4}$$

measures the relative change in its marginal cost that a firm achieves by moving the routine task abroad. Assuming that offshoring also entails a fixed cost resulting from the purchase of offshoring services, it is only attractive for source country producers to move routine tasks abroad if $\kappa > 1$, making κ the *marginal cost savings factor* that a firm can achieve by offshoring. While κ is endogenous and yet to be determined, it is immediate that the equilibrium will feature offshoring, provided that variable offshoring costs τ are finite: If no firm were to offshore, w^* would fall to zero since the host country has no local entrepreneurs. Eq. (3.4) shows that in this case $c^o(v)$ would fall to zero, which implies that at least some firms would self-select into offshoring.

Firms set prices as a constant markup $1/\rho$ over marginal cost, giving

$$p^i(v) = \frac{c^i(v)}{\rho} \qquad i \in \{d, o\}. \tag{3.5}$$

Using Eqs. (3.2), (3.4) and (3.5), we can compute relative operating profits of two firms with the same productivity, but differing offshoring status. We get:[11]

$$\frac{\pi^o(\varphi)}{\pi^d(\varphi)} = \kappa^{\sigma-1}. \tag{3.6}$$

With $\kappa > 1$, an offshoring firm makes higher operating profits than a purely domestic firm with identical productivity. Analogously, the relative operating profits by two firms with the same offshoring status but differing productivities φ_1 and φ_2 are given by

$$\frac{\pi^i(\varphi_1)}{\pi^i(\varphi_2)} = \left(\frac{\varphi_1}{\varphi_2}\right)^{\xi}, \qquad i \in \{d, o\}, \tag{3.7}$$

where $\xi \equiv \sigma - 1$. Therefore, given their offshoring status, more productive firms make higher operating profits.

[10] We use an asterisk to denote variables pertaining to the host country of offshoring.

[11] We suppress firm index v from now on, because a firm's performance is fully characterised by its position in the productivity distribution and its offshoring status.

3.1.3 Equilibrium factor allocation

We assume that the source and the host country of offshoring are populated by N and N^* agents, respectively. While the population in the host country has only access to a single activity, namely the performance of routine tasks in the foreign affiliates of offshoring firms, agents in the source country can choose from a set of three possible occupations: entrepreneurship, employment as a production worker, and employment in the offshoring-service sector.[12] Entrepreneurs are owner-managers of firms, and their ability determines firm productivity. To keep things simple, we assume that entrepreneurial ability maps one-to-one into firm productivity, and we can therefore use a single variable, φ, to refer to ability as well as productivity. Being the residual claimant, the entrepreneur receives firm profits as individual income. Agents differ in their entrepreneurial abilities, and hence in the profits they can achieve when running a firm. Following standard practice, we assume that abilities (and thus productivities) follow a Pareto distribution, for which the lower bound is normalised to one: $G(\varphi) = 1 - \varphi^{-k}$, and where both $k > 1$ and $k > \xi$ are assumed in order to guarantee that the mean of firm-level productivities and the mean of firm-level revenues, respectively, are positive and finite.

Entrepreneurial ability is irrelevant for the two alternative activities that can be performed in the source country of offshoring, so that agents are symmetric in this respect. If an individual works in the offshoring-service sector, she receives a fee s, which is determined in a perfectly competitive market in general equilibrium. Finally, agents in the source country can also apply for a job as production worker and perform the routine or non-routine task, receiving the endogenous wage rate w. As shown below, our equilibrium features self-selection of the most productive firms into offshoring if the variable cost of offshoring is sufficiently high. In this case, the lowest-productivity firm is purely domestic. Denoting this firm's productivity by φ^d, we can characterize the marginal entrepreneur by indifference condition

$$\pi^d(\varphi^d) = w = s. \tag{3.8}$$

[12] It is not essential for our analysis that source country labour is used for providing offshoring services. This assumption mediates factor reallocations between entrepreneurship and employment as production workers – which are essential for the main results in this chapter – and hence it helps us to secure against overemphasizing the role of occupational changes in the source country of offshoring.

We assume that offshoring requires the purchase of one unit of offshoring services and that the labour input coefficient in the service sector is equal to one.[13] The indifference condition for the entrepreneur running the marginal offshoring firm with productivity φ^o is given by

$$\pi^o(\varphi^o) - \pi^d(\varphi^o) = s, \tag{3.9}$$

i.e. for the indifferent entrepreneur the gain in operating profits achieved by offshoring equals the fixed offshoring cost. All variables in Eqs. (3.8) and (3.9) are endogenous, and both indifference conditions are linked via their dependence on s. To illustrate the nature of this link, consider some change in the value of model parameters that leads to, say, an increase in w. As a consequence, the fee s paid to individuals in the offshoring service sector has to increase by the same amount in order to keep individuals indifferent between both occupations. A higher offshoring service fee s drives up the fixed cost of offshoring, thereby in turn requiring a larger offshoring-induced gain in operating profits in order to keep the marginal offshoring firm indifferent between both modes of operation. We now proceed in two steps: in the remainder of this section we solve for the domestic factor allocation as a function of model parameters and the fraction of offshoring firms $\chi \equiv [1 - G(\varphi^o)]/[1 - G(\varphi^d)]$, while in Section 3.1.4 below we link χ to the underlying model parameters, including the (variable) costs of offshoring τ.

The indifference condition in Eq. (3.8) postulates the equality between profits of the marginal firm $\pi^d(\varphi^d)$ and the wage rate of production workers w. We now link these two variables to economy-wide aggregates. For this purpose, it is useful to introduce three new operating profit averages, namely average operating profits $\bar{\pi}$, average operating profits for the counterfactual situation in which all firms would choose domestic production $\bar{\pi}^{\text{dom}}$ and the average operating profit surplus due to the most productive firms actually choosing offshoring instead of domestic production $\bar{\pi}^{\text{off}}$. There is a direct relation between the three averages which is given by $\bar{\pi} = \bar{\pi}^{\text{dom}} + \chi\bar{\pi}^{\text{off}}$. Due to Pareto distributed productivities, the two averages $\bar{\pi}^{\text{dom}}$ and $\bar{\pi}^{\text{off}}$ are linked to operating profits of the marginal domestic firm $\pi^d(\varphi^d)$ and the gain in operating profits of the marginal offshoring firm $\pi^{\text{off}}(\varphi^o) \equiv \pi^o(\varphi^o) - \pi^d(\varphi^o)$, respectively, by the factor of proportionality

[13] Our analysis extends in a straightforward way to the more general case where firms require $f^o > 0$ units of offshoring services.

$\zeta \equiv k/(k-\xi)$. This allows us to write

$$\bar{\pi} = \zeta \left[\pi^d(\varphi^d) + \chi\pi^{\text{off}}(\varphi^o)\right] = \zeta(1+\chi)\pi^d(\varphi^d),$$

where the second equality follows from the fact that due to indifference conditions (3.8) and (3.9) both $\pi^d(\varphi^d)$ and $\pi^{\text{off}}(\varphi^o)$ are equal to s. Using the relation $\sigma\bar{\pi} = Y/M$, we can express profits of the marginal firm as a function of economy-wide variables:

$$\pi^d(\varphi^d) = \frac{1}{\zeta}\frac{Y}{\sigma M(1+\chi)}. \tag{3.10}$$

Turning to the determination of w, we make use of the fact that due to constant markup pricing the wage bill of each source country firm is a constant fraction ρ of the firm's revenues. Taking into account the fact that for offshoring firms only a fraction η of the wage bill is paid to production workers in the source country, and denoting by $\bar{\pi}^d$ and $\bar{\pi}^o$ the average operating profits of purely domestic and offshoring firms, respectively, we get

$$w = \gamma\rho\frac{Y}{L}, \tag{3.11}$$

where

$$\gamma \equiv \frac{(1-\chi)\bar{\pi}^d + \chi\eta\bar{\pi}^o}{\bar{\pi}}$$

is the share of the overall wage bill paid in the source county, and L is the endogenous supply of source country production workers.[14] We show in Appendix A.2 that γ can be written as

$$\gamma(\chi;\eta) = \frac{1+\eta\chi-(1-\eta)\chi^{\frac{k-\xi}{k}}}{1+\chi}.$$

It is easily confirmed that $\gamma(\chi;\eta)$ decreases monotonically in χ, falling from the maximum value of 1 at $\chi = 0$ to the minimum value of η at $\chi = 1$.

Having derived, in Eqs. (3.10) and (3.11), expressions for the wage rate of production workers and the profit income of the marginal entrepreneur, respectively, we can rewrite indifference condition (3.8) as:

$$L = \gamma\zeta\,(1+\chi)\,(\sigma-1)M. \tag{3.12}$$

[14]To simplify notation, we suppress the arguments of functions when the dependence is clear from the context.

A second condition linking L and M is established by the resource constraint

$$L = N - (1 + \chi) M, \tag{3.13}$$

which illustrates that individuals can work as either entrepreneurs (M), workers in the service sector (χM), or production workers (L). Together, Eqs. (3.12) and (3.13) pin down the equilibrium mass of intermediate goods producers M and the equilibrium mass of production workers L as functions of model parameters and a single endogenous variable, the share of exporting firms χ:

$$M = \left\{ \frac{1}{(1+\chi)\left[1 + \gamma\zeta(\sigma - 1)\right]} \right\} N, \tag{3.14}$$

$$L = \left[\frac{\gamma\zeta(\sigma - 1)}{1 + \gamma\zeta(\sigma - 1)} \right] N. \tag{3.15}$$

The mass of firms is linked to the ability of the marginal entrepreneur by the condition $M = [1 - G(\varphi^d)]N$, and solving for φ^d gives

$$\varphi^d = \left\{ (1+\chi)\left[1 + \gamma\zeta(\sigma - 1)\right] \right\}^{\frac{1}{k}}. \tag{3.16}$$

In the next subsection we show how χ is determined as a function of the cost of offshoring τ.

3.1.4 Determining the share of offshoring firms

In this subsection, we derive the formal condition in terms of model parameters for an interior offshoring equilibrium, i.e. a situation in which some but not all firms offshore, and we also show how the share of offshoring firms χ varies with the cost of offshoring τ in an interior equilibrium.

Given our assumption of Pareto distributed productivities, the indifference condition of the marginal offshoring firm (3.9) allows us to derive a link between χ and the marginal cost savings factor κ. Substituting from Eqs. (3.6) to (3.8), we get the *offshoring indifference condition* (OC)

$$\chi = \frac{1 - G(\varphi^o)}{1 - G(\varphi^d)} = \left(\kappa^{\sigma-1} - 1\right)^{\frac{k}{\xi}}. \tag{3.17}$$

Intuitively, a larger marginal cost savings factor κ makes offshoring more attractive, and therefore a larger share of firms chooses to move production of their routine tasks abroad. It is easily checked in Eq. (3.17) that an interior equilibrium with $\chi \in (0, 1)$ requires $1 < \kappa < 2^{1/(\sigma-1)}$.

A second link between χ and κ can be derived from the condition for labour market equilibrium in both countries. Labour market equilibrium in the source country follows from Eqs. (3.11) and (3.15) as

$$w = \rho \left[\frac{1 + \gamma\zeta(\sigma - 1)}{\zeta(\sigma - 1)} \right] \left(\frac{Y}{N} \right),$$

while labour market equilibrium in the host country is analogously given by

$$w^* = (1 - \gamma)\, \rho \left(\frac{Y}{N^*} \right).$$

Using Eq. (3.4), we arrive at the *labour market constraint* (LC), which links labour market equilibrium in both countries to the marginal cost savings factor κ:

$$\kappa = \left[\frac{1 + \gamma\zeta(\sigma - 1)}{\tau\, (1 - \gamma)\, \zeta(\sigma - 1)} \left(\frac{N^*}{N} \right) \right]^{1-\eta}. \tag{3.18}$$

Since γ decreases monotonically from 1 to η as χ increases from zero to one, we know that the labour market constraint is monotonically decreasing in χ, starting from infinity. This is intuitively plausible: At $\chi = 0$, there is no production in the host country, and wage rates there fall to zero, making the marginal cost savings factor κ infinitely large. Holding τ constant, as more firms start to offshore production, effective wages in the host country are bid up, thereby reducing κ.

Combining Eqs. (3.17) and (3.18) in Figure 3.1, we can conclude that an interior equilibrium with $\chi < 1$ is reached if the right-hand side of Eq. (3.18), evaluated at $\gamma(1, \eta) = \eta$, is smaller than $2^{1/(\sigma-1)}$. This can obviously be achieved for sufficiently high values of τ, because a higher τ lowers for any given χ the marginal cost-saving factor of offshoring determined by the right-hand side of Eq. (3.18), while leaving the link between χ and κ established by the offshoring indifference condition in Eq. (3.17) unaffected. A decline in the relative population size N^*/N has a similar effect. The smaller the relative size of the host country population, the larger is, all other things equal, the endogenous relative wage $\tau w^*/w$, and hence the smaller are the potential cost savings from offshoring, according to Eq. (3.18). Therefore, focusing on an interior equilibrium with $\chi \in (0, 1)$ is equivalent to focusing on sufficiently high levels of τ and/or sufficiently low levels of N^*/N, and this is what we do in the subsequent analysis. Such an interior equilibrium is illustrated in Figure 3.1.

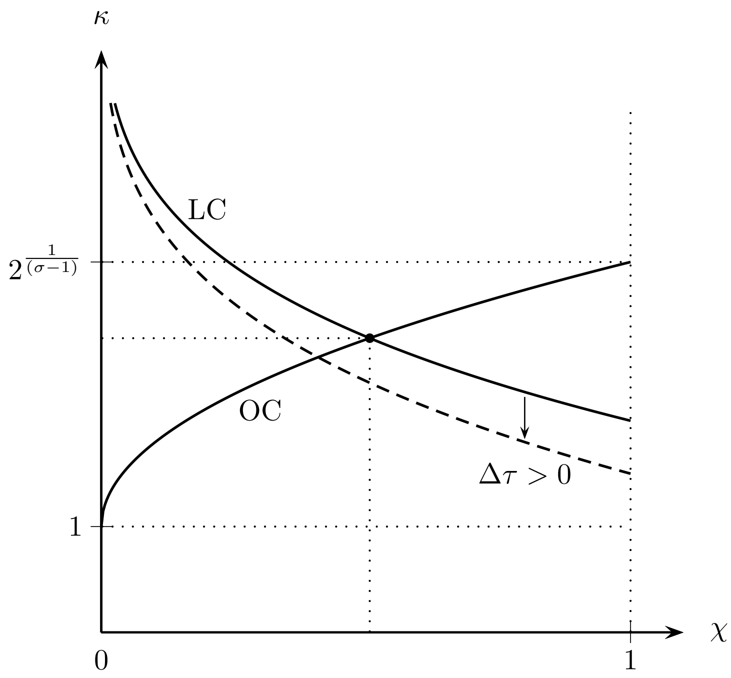


Figure 3.1: *Partitioning of firms by their offshoring status*

To get insights on the link between offshoring cost τ and the share of offshoring firms χ, we can combine Eqs. (3.17) and (3.18) to the implicit function

$$F(\chi, \tau) \equiv \left[\frac{1+\gamma\zeta(\sigma-1)}{\tau\,(1-\gamma)\,\zeta(\sigma-1)}\left(\frac{N^*}{N}\right)\right]^{1-\eta} - \left(1+\chi^{\frac{\xi}{k}}\right)^{\frac{1}{\sigma-1}} = 0.$$

Implicit differentiation yields $d\chi/d\tau < 0$ for any interior equilibrium with $0 < \chi < 1$. As noted above, higher direct costs of shipping intermediate goods, i.e. a higher parameter τ, shifts the LC locus downwards, but does not affect the OC locus in Figure 3.1. We therefore have the intuitive result that a higher τ reduces the marginal cost savings factor κ, and thus reduces χ, the equilibrium share of firms that shift production of their routine task abroad. Due to the monotonic relationship between (endogenous) χ and (exogenous) τ we can equivalently derive comparative static results below in terms of either variable.[15]

[15]One can see in Eq. (3.18) that the limiting case $\chi \to 0$ is induced by $\tau \to \infty$.

3.2 The effects of offshoring

The purpose of this section is to look at the effects of offshoring on key economic variables, namely on the factor allocation between occupations and between firms, on income inequality within the group of entrepreneurs as well as between entrepreneurs and production workers, and on aggregate welfare. Throughout this section, we derive comparative static results in terms of changes in χ. As shown above, this is equivalent to considering exogenous changes in offshoring cost τ, noting that $d\chi/d\tau < 0$, and hence in the discussion of results we will sometimes refer to changes in τ as well. Also, we focus our discussion on the source country, since the effects for the host country are trivial due to our simplifying assumption that no firms are headquartered there.

3.2.1 Factor allocation

Since our economy is populated by firms of well-defined size, we can distinguish between allocation effects at the firm level and economy-wide allocation effects. Looking first at the firm level, we ask the question what the offshoring decision does to employment of a firm in the source country. Firm-level employment in the source country for an offshoring firm and for a purely domestic firm, respectively, follows from applying Shephard's Lemma to the firm-specific variable unit cost functions in Eq. (3.4), and multiplying the resulting labour input coefficients by firm-level output. This gives:

$$l^o(\varphi) = \frac{\eta q^o(\varphi)}{\varphi\kappa} \qquad \text{and} \qquad l^d(\varphi) = \frac{q^d(\varphi)}{\varphi},$$

respectively. The source-country employment effect of offshoring at the firm level can now be computed as the log difference $\ln l^o(\varphi) - \ln l^d(\varphi)$, which is the difference in percent between domestic employment of an offshoring firm and employment of a purely domestic firm with the same productivity. The firm-level employment effect thus measured compares for each firm the actual employment level with the employment in a counterfactual situation in which the respective firm would be in the other category.

To get a better intuition, it is helpful to write the firm-level effect as the sum of two partial effects, the effect of offshoring on employment per unit of output, and the effect of offshoring on firm-level output. We call the first effect the *international relocation effect* (IR), since it measures the direct effect of relocating tasks abroad on firm-level employment in the source country, without taking into account the induced reduction in marginal cost. The second effect we call the *firm-level productivity effect* (FP), since it is a measure of the change in output – and, hence, the change in employment – induced by the reduction in marginal cost.[16] Using the link between κ and χ given in offshoring indifference condition Eq. (3.17), we obtain

$$\ln l^o(\varphi) - \ln l^d(\varphi) = \underbrace{\ln\left[\eta\left(1+\chi^{\frac{\xi}{k}}\right)^{\frac{1}{1-\sigma}}\right]}_{\text{IR}} + \underbrace{\ln\left[\left(1+\chi^{\frac{\xi}{k}}\right)^{\frac{\sigma}{\sigma-1}}\right]}_{\text{FP}}. \tag{3.19}$$

The international relocation effect is negative for any $\chi \geq 0$, since on the one hand the routine task is now produced by foreign labour and on the other hand the input ratio changes in favour of this – now relatively cheaper – task. The latter effect is stronger if the marginal cost savings factor κ is higher, i.e. if χ is higher. In contrast to the international relocation effect, the firm-level productivity effect is zero if evaluated at $\chi = 0$ (since the marginal cost savings factor κ is zero), and it increases monotonically with increasing κ, i.e. with increasing χ.

Two aspects of the partial firm-level employment effects identified above are noteworthy. First, Eq. (3.19) shows that neither effect depends on firm productivity. Hence, for a given level of offshoring costs, implying some value of χ, the percentage difference in firm-level domestic employment relative to the respective counterfactual (offshoring for the purely domestic firms, purely domestic production for the offshoring firms) is the same for all firms. Second, the fact that only the international relocation effect is of first order at $\chi = 0$, while both effects are continuous in χ, means that the international relocation effect determines the overall effect at low levels of offshoring. Inspection of Eq. (3.19) furthermore shows that the firm-level productivity effect dominates at high levels of offshoring if and only if the cost share of non-routine tasks η is greater than 0.5. This is the case we focus on in the following, which is in line with the findings of Blinder (2009) and Blinder and Krueger (2013), who report for the US that 25 percent of tasks can be classified as offshorable and thus could be moved abroad in principle. While this

[16]The effects are directly analogous to the labour supply effect and the productivity effect, respectively, derived by Grossman and Rossi-Hansberg (2008), but in contrast to the latter they are identified at the firm level rather than just at the aggregate level.

number is not a perfect match for our cost-share parameter η, the fact that the Blinder-Krueger measure considers potential offshorability rather than actual offshoring renders our parameter constraint of $\eta > 0.5$ a rather conservative assumption.[17]

The firm-level employment effects of the decentralised offshoring decisions have consequences for the allocation of domestic workers across firms. Considering a decrease in marginal costs of offshoring τ, Eq. (3.19) describes the effect on the employment in marginal (newly) offshoring firms, which is negative at high levels of τ and positive if τ is low. To derive the effect on the employment in infra-marginal firms (purely domestic firms and incumbent offshoring firms) we use the result that due to constant-markup pricing relative employment across firms in the same category is identical to relative operating profits, and therefore in analogy to Eq. (3.7) given by $l^i(\varphi_1)/l^i(\varphi_2) = (\varphi_1/\varphi_2)^\xi$. In addition, also as a consequence of constant-markup pricing, the wage bill of the marginal firm is a multiple $\sigma - 1$ of its operating profits, and with $w = \pi^d(\varphi^d)$ we find that employment of the marginal firm is given by $l^d(\varphi^d) = \sigma - 1$.

Using these results, Figure 3.2 illustrates the effects of a decrease in τ on the allocation of production labour, where the top panel shows the case of low χ (high τ), while the bottom panel shows the case of high χ (low τ). If χ is low, a marginal reduction in τ increases employment in all purely domestic firms (of which there are relatively many), including – as shown formally below – some new entrants. It also increases employment in the incumbent offshoring firms (of which there are relatively few). The newly offshoring firms – which are high-productivity firms in this case – are therefore the only ones to shed production workers in the source country if τ is reduced and the share of offshoring firms is low. If χ is high the picture is different: following a decrease in τ employment in all offshoring firms, marginal and infra-marginal, increases, while employment in purely domestic firms falls, and the least productive firms stop production and exit. Hence, offshoring exerts a non-monotonic effect on the allocation of production workers

[17] Empirical evidence for the effect of offshoring on firm-level employment comes from Moser, Urban, and Weder di Mauro (2009), Hummels, Jørgensen, Munch, and Xiang (2013) and Monarch, Park, and Sivadasan (2013), who sort out the firm-level productivity effect and the international relocation effect using matched employer-employee-data. While the former study finds that the firm-level productivity effect dominates for the case of Germany, the opposite seems to occur in Denmark and the US as noted by Hummels, Jørgensen, Munch, and Xiang (2013) and Monarch, Park, and Sivadasan (2013), respectively.

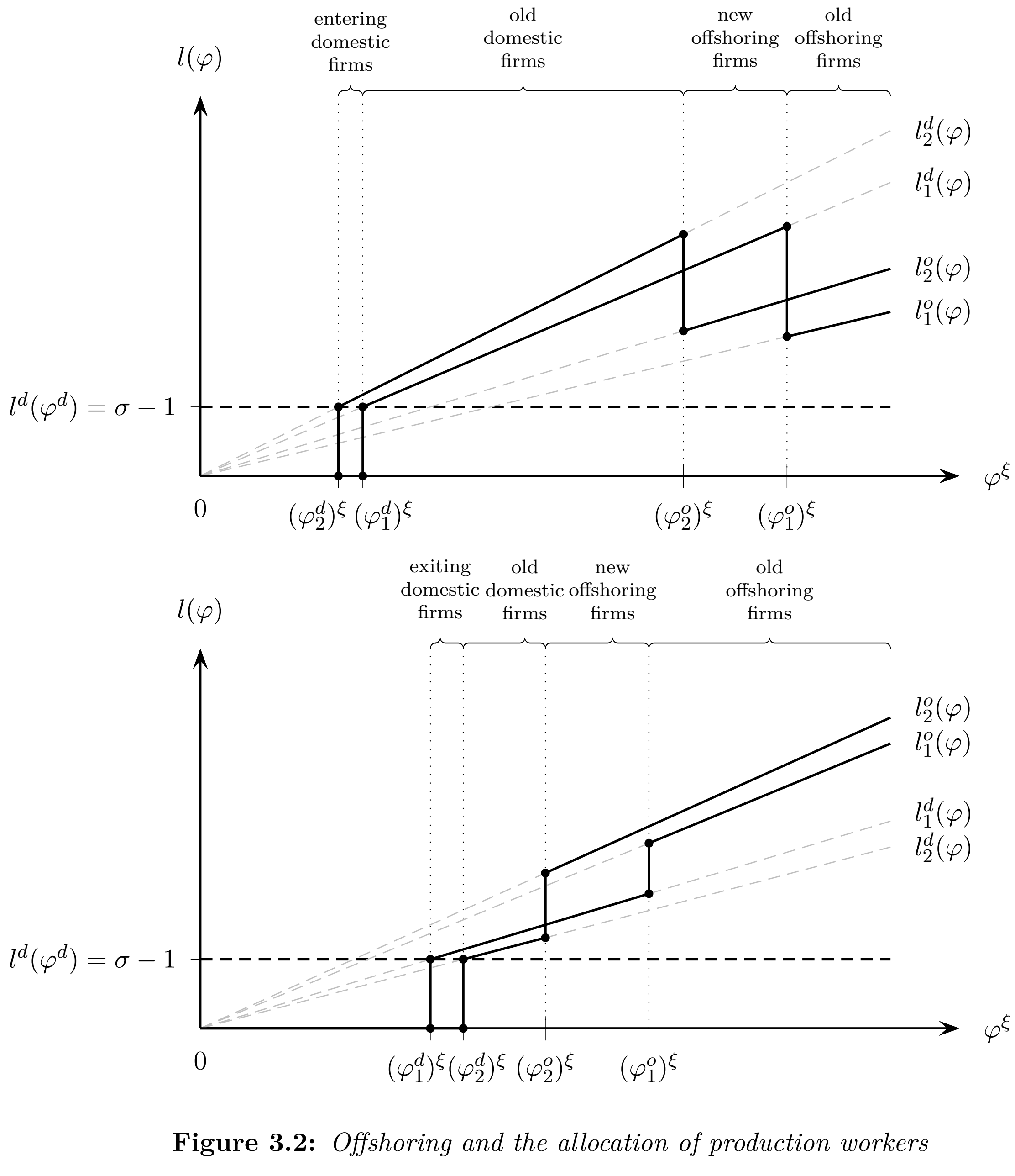


Figure 3.2: *Offshoring and the allocation of production workers*

across firms, reallocating them towards less productive firms if offshoring costs are high, and towards more productive ones if offshoring costs are low.

The effect of offshoring on aggregate factor allocation in our model works via its effect on occupational choice, considering that the labour indifference condition has to hold throughout. Formally, the effects of offshoring on the mass of production workers and the mass of firms follow directly from Eqs. (3.14) and (3.15):

$$\frac{dL}{d\chi} = \frac{\zeta(\sigma-1)\partial\gamma/\partial\chi}{[1+\gamma\zeta(\sigma-1)]^2}N, \quad \frac{dM}{d\chi} = -\frac{1+\zeta(\sigma-1)\left[\gamma+(1+\chi)\partial\gamma/\partial\chi\right]}{(1+\chi)^2[1+\gamma\zeta(\sigma-1)]^2}N.$$

Since $\partial\gamma/\partial\chi$ is negative, it is immediate that $dL/d\chi < 0$ holds for arbitrary levels of χ, and hence in line with the empirical findings of Ebenstein, Harrison, McMillan, and Phillips (2013) offshoring unambiguously reduces the mass of production workers in our model, with the affected individuals either moving to the offshoring service sector, or becoming managers of newly-opened low-productivity firms.

The effect of offshoring on the mass of firms (or, equivalently, on the cutoff productivity of the marginal firm) is non-monotonic, with $dM/d\chi > 0$ for low levels of χ and $dM/d\chi < 0$ when χ is high. If χ is close to zero and τ is reduced, the newly offshoring firms are the most productive ones and these are the firms with the largest workforce in both tasks. Not all workers losing their jobs in these firms can be absorbed by expansion of other already existing firms or by expansion of the offshoring service sector, and hence new firms have to enter in order to restore the labor market equilibrium. For low levels of χ, M therefore increases as τ decreases. The effects are different for high levels of χ, because labour demand from offshoring firms (new and old) increases as τ decreases, and the mass of firms has to fall in order to restore the labour market equilibrium.[18] The effects are summarized in the following proposition:

Proposition 3.2.1 *When χ is low, a reduction in marginal offshoring costs τ reallocates production workers towards less productive firms, and new firms enter the market in the lower tail of the productivity distribution. When χ is high, a reduction in τ reallocates production workers towards more productive firms, and firms at the lower tail of the productivity distribution leave*

[18] To see these effects formally, consider $\eta > 0.5$ and note that $\partial\gamma/\partial\chi$ is equal to $-\infty$ if evaluated at $\chi = 0$ and equal to $(\eta-1)/(2\zeta)$ if evaluated at $\chi = 1$.

the market. The mass of production workers decreases monotonically with a decrease in τ*.*

Proof Analysis in the text.

The potentially unfavourable effect of offshoring on the resource allocation in the source country constitutes a key difference to international trade in goods, which in a comparable setting always reallocates labour from low- to high-productivity firms (cf. Egger and Kreickemeier, 2012), with the latter effect of course well known from the canonical model by Melitz (2003). The finding that offshoring in our setting has a non-monotonic effect on labour allocation is furthermore a direct consequence of firm heterogeneity. To see this, consider the limiting case of $k \to \infty$, in which all firms have the same productivity (equal to 1, the lower bound of the Pareto distribution). In this model variant, both the international relocation effect and the firm-level productivity effect are independent of the level of χ and, according to Eq. (3.19), they are given by $\ln[\eta 2^{1/(1-\sigma)}]$ and $\ln[2^{\sigma/(\sigma-1)}]$, respectively. Consequently, the firm-level productivity effect of offshoring is of first order already at $\chi = 0$, whereas the adverse international relocation effect is mitigated, because the newly offshoring firms have lower employment than in the model variant with heterogeneous producers. A reduction in τ therefore reallocates production workers towards offshoring firms, and some firms leave the market for any $\chi \in (0, 1)$.

3.2.2 Inequality among entrepreneurs and between groups

Intra-group inequality of entrepreneurial income is measured by the Gini coefficient for profit income, which, as formally shown in Appendix A.3, is given by

$$A_M(\chi) = \frac{\zeta - 1}{\zeta + 1}\left[1 + \frac{\chi\,(2-\chi)}{\zeta + (\zeta - 1)\chi}\right]. \tag{3.20}$$

The relationship between Gini coefficient $A_M(\chi)$ and the share of offshoring firms χ is non-monotonic as offshoring always increases the profits of newly offshoring firms. If the share of offshoring firms is small, an increase in χ implies that newly offshoring firms are run by entrepreneurs with high ability, and these are firms that already ranked high in the profit distribution prior to offshoring. Hence, an increase in χ raises the dispersion of profit income in this case. Things are different at high levels of χ, because newly offshoring firms are now firms with a low rank in the distribution of profit income and an increase in χ therefore lowers the dispersion of profit income. Furthermore, comparing $A_M(\chi)$ for $\chi > 0$ with $A_M(0)$, we find that offshoring increases the dispersion of profit income relative to the benchmark without offshoring, irrespective of the prevailing level of χ. This result is due to the fact that the common fixed cost of offshoring disproportionately affects the profits of less productive firms, thereby contributing to an increase in the dispersion of profit incomes.[19]

Inter-group inequality is measured by the ratio of average entrepreneurial income and average labour income, where the latter is simply given by wage rate w. According to Eq. (3.10), average entrepreneurial income, $\bar{\psi}$, is equal to $\pi^d(\varphi^d)(1+\chi)\zeta - \chi s$. Applying indifference condition (3.8), the ratio of average entrepreneurial income and average income of production workers, $\bar{\omega} \equiv \bar{\psi}/w$, is therefore given by

$$\bar{\omega} = \zeta + (\zeta - 1)\chi. \tag{3.21}$$

It follows immediately from Eq. (3.21) that inter-group inequality rises monotonically in the share of offshoring firms χ. The intuition is as follows: A higher value of χ indicates that the marginal cost saving factor κ must be higher, which in turn implies that profits of all offshoring firms increase, both in absolute terms and relative to the profits of the marginal firm in the market. Since the marginal firm's profits are equal to w, it is clear that inter-group inequality has to go up in response to an increase in χ.

The following proposition summarises the results.

Proposition 3.2.2 *The inequality of entrepreneurial income, measured by the Gini coefficient, rises with the share of offshoring firms at low levels of χ, and decreases at high levels of χ, while*

[19] Since the offshoring service sector is perfectly competitive, one can think of individuals working there as one-person firms, and hence we can define the group of self-employed agents, which covers both entrepreneurs and offshoring service providers. As we show in Appendix A.4, the Gini coefficient for this broadly defined income group can be expressed as

$$A_S(\chi) = \frac{\zeta - 1}{\zeta + 1}\left[1 + \frac{2}{\zeta}\frac{\chi}{(1+\chi)^2}\right],$$

with $A_S'(\chi) > 0$. Therefore, inequality in the group of all self-employed agents increases monotonically with χ. The comparison of $A_M(\chi)$ and $A_S(\chi)$ furthermore shows that inequality within the group of all self-employed agents is less pronounced than inequality within the subgroup of entrepreneurs.

always staying higher than in the benchmark situation without offshoring. Increasing the share of offshoring firms χ leads to a monotonic increase in inter-group inequality between entrepreneurs and workers.

Proof Analysis in the text and formal discussion in Appendix A.3.

Together the two effects in Proposition 3.2.2 give birth to a class of entrepreneurial superstars (cf. Gabaix and Landier, 2008), who benefit from the global expansion of their respective firms by sourcing part of their production from low-cost locations abroad.

3.2.3 Welfare

With just a single global consumption good, welfare for the source country is simply given by source country income per capita. Aggregate income in the source country is given by $I = (1-\rho+\gamma\rho)Y$, where $(1-\rho)Y$ is the sum of profit income and offshoring service income, and $\gamma\rho Y$ is domestic labour income. The determination of the welfare effects of offshoring is the only place in our analysis where the extent of external increasing returns to scale, introduced earlier in Eq. (3.1) via parameter $\varepsilon \in [0,1]$, matters for the results. Using Eq. (3.2) for the marginal firm with productivity φ^d, as well as Eqs. (3.8) and (3.10), we get

$$I(\chi) = (1-\rho+\gamma\rho)\,\mathcal{A}\,(1+\chi)^{\frac{\sigma}{\sigma-1}}\left(M\varphi^d\right)M^{\frac{\varepsilon}{\sigma-1}}, \tag{3.22}$$

with $\mathcal{A} \equiv (\sigma-1)\zeta^{\sigma/(\sigma-1)}$ collecting parameters, and a solution for I in terms of model parameters and χ follows by substituting for M and φ^d from Eqs. (3.14) and (3.16), respectively. Income in the source country is higher in an offshoring equilibrium than in autarky if for the specific share χ of offshoring firms in this equilibrium we have $\Phi(\chi) \equiv I(\chi)/I(0) > 1$, and lower than in autarky if $\Phi(\chi) < 1$. It is easy to see that ε plays a crucial role for the welfare effect of offshoring: the greater the external increasing returns to scale, the more beneficial is an increase in the mass of produced varieties M for aggregate output, ceteris paribus, and therefore the less harmful will be the resource allocation towards less productive firms that, as shown above, is characteristic for offshoring at low levels of χ.

For the sake of transparency, we start with the discussion of the two polar cases $\varepsilon = 0$ and $\varepsilon = 1$. If $\varepsilon = 0$, there are no external increasing returns to scale, and the mass of firms has no independent effect on aggregate output. As we show formally in Appendix A.5, source country welfare in this case is lower than in autarky if the level of offshoring is low, and it is higher than in autarky if the level of offshoring is high. The sign of the welfare effect is determined by two partial effects: an expansion of economy-wide output, that can be achieved by using foreign labour to perform routine tasks at lower cost, and an outflow of labour income because foreign workers must be paid by offshoring firms. The relative strength of these two counteracting effects depends on the relative strength of international relocation and firm-level productivity effect. Therefore, the main forces determining the welfare implications of offshoring are the same as the forces determining its implications for labour allocation. Offshoring reallocates labour towards less productive uses if χ is low, and in this case source country welfare falls. By contrast, offshoring reallocates labour towards more productive uses if χ is high, and in this case source country welfare increases.

The reallocation effect is of course also welfare relevant if $\varepsilon = 1$, and viewed on its own it leads to a welfare decrease at low levels of χ. But the increase in the mass of varieties now affects welfare positively, and hence overall offshoring has an unambiguously positive effect on welfare. Intuitively, this is so since with $\varepsilon = 1$ decentralized entry decisions establish allocational efficiency, and hence the market outcome replicates the solution to the social planner's problem in the source country under autarky (see Appendix A.6). Offshoring provides access to (cheap) foreign labour, and this expands domestic production possibilities with positive welfare implications for the source country, as in Grossman and Rossi-Hansberg (2008). We show in Appendix A.5 that the welfare results for the two borderline cases carry over to intermediate cases of ε. In particular, we derive a critical value $\bar{\varepsilon} \equiv (\sigma - \xi)(\sigma - 1)/(\sigma k)$ and show that offshoring is detrimental for source-country welfare at low levels of χ if the external increasing returns to scale are sufficiently weak ($\varepsilon < \bar{\varepsilon}$), while offshoring is always beneficial for the source country if the external increasing returns to scale are sufficiently strong ($\varepsilon > \bar{\varepsilon}$).[20]

Taking stock, source country welfare in our model can only fall as a consequence of offshoring if the factor allocation is not efficient under autarky, and hence $\varepsilon < 1$ is a necessary condition for

[20] In Appendix A.5, we also show that the external increasing returns to scale reported by Ardelean (2011) for the US are sufficiently small to render the welfare losses of offshoring at low levels of χ empirically relevant.

welfare losses. In this case, offshoring can lower source country welfare by reallocating workers towards less productive uses. Domestic misallocation of resources as a potential source of losses from offshoring in the case $\varepsilon < 1$ distinguishes our model from similar results in Grossman and Rossi-Hansberg (2008) and Rodriguez-Clare (2010), where source country welfare can fall due to an offshoring-induced negative terms-of-trade effect. As outlined above, offshoring cannot have such unfavourable allocation effects in our model if firms are identical, and hence welfare losses from offshoring in our model are the result of a misallocation of resources in the presence of heterogeneous firms. This relates our analysis to Dhingra and Morrow (2013) who construct a model with monopolistic competition among heterogeneous firms in which endogenous markups lead to a misallocation of resources that can be amplified by trade.

We summarise our insights regarding the welfare implications of offshoring in the following proposition.

Proposition 3.2.3 *For strong external increasing returns to scale welfare in the source country increases monotonically in the share of offshoring firms. For weak external increasing returns to scale welfare in the source country decreases with the share of offshoring firms at low levels of χ. The effect is reversed as more firms offshore, and welfare surpasses its autarky level if χ is sufficiently large.*

Proof Analysis in the text and formal discussion in Appendix A.5.

3.3 Offshoring in the presence of firm-level rent-sharing

In this section, we extend our framework by a more sophisticated labour market model, which allows us to address the widespread concern that offshoring may have a negative effect on aggregate employment in a country that shifts production of routine tasks to a low-wage location (cf. Geishecker, Riedl, and Frijters, 2012). More specifically, we develop a model of firm-level rent sharing that features involuntary unemployment of production workers and, at the same time, captures the stylised fact that more profitable firms pay higher wages (cf. Blanchflower, Oswald, and Sanfey, 1996).[21]

The labour market model proposed in this section is a fair-wage-effort model which builds upon the idea of gift exchange, and whose main assumptions are rooted in insights from psychological research (see Akerlof, 1982; Akerlof and Yellen, 1990). The model postulates a positive link between a firm's wage payment and a worker's effort provision, and workers exert full effort, normalised to equal one, if and only if they are paid at least the wage they consider fair.[22] As in Egger and Kreickemeier (2012) and Egger, Egger, and Kreickemeier (2013) we assume that the fair wage $\hat{w}$ is a weighted average of firm-level operating profits $\pi(\varphi)$ and the average wage of production workers $(1-U)\bar{w}$, where U is the unemployment rate of production workers and $\bar{w}$ is the average wage of those production workers who are employed:

$$\hat{w}(\varphi) = [\pi(\varphi)]^{\theta}[(1-U)\bar{w}]^{1-\theta}, \qquad \theta \in (0,1). \tag{3.23}$$

An analogous condition, with $(1-U^*)\bar{w}^*$ substituted for $(1-U)\bar{w}$, holds in the host country of offshoring, which implies that multinationals share their rents with workers in the source *and* host country of offshoring.[23] Following Akerlof and Yellen (1990), we assume that effort decreases proportionally with the wage if workers are paid less than $\hat{w}$, and hence firms have no incentive to do so. At the same time, as we discuss below, our model features involuntary unemployment in equilibrium, and therefore even low-productivity firms do not need to pay more than $\hat{w}$ to attract workers. Firms hence set $w(\varphi) = \hat{w}(\varphi)$, and Eq. (3.23) describes the distribution of wages across firms as a function of firm-level operating profits.[24]

[21]Offshoring in the presence of labour market imperfections is also discussed in other papers, including Egger and Kreickemeier (2008), Keuschnigg and Ribi (2009) and Mitra and Ranjan (2010). While all of these studies highlight important channels through which offshoring can affect domestic employment, neither study sheds light on the specific role of firm heterogeneity or the consequences of occupational choice.

[22]Fehr, Goette, and Zehnder (2009) survey the extensive experimental evidence for the fair-wage-effort hypothesis. Cohn, Fehr, and Goette (2013) provide evidence supportive of the fair-wage-effort hypothesis in a field study.

[23]Evidence supportive of international rent sharing within firms is provided by Budd, Konings, and Slaughter (2005) and Martins and Yang (2013).

[24]Even though firms set wages unilaterally, their profit maximisation problem does not differ from the one in Section 3.1.2. As pointed out by Amiti and Davis (2012), wages depend positively on profits due to fair-wage

In contrast to the full employment version of our model the decision to become a production worker in a labour market with firm-specific wages now carries an income risk.[25] We make the standard assumption that workers have to make their career choice before they know the outcome of the job allocation process among applicants (cf. Helpman and Itskhoki, 2010).[26] With risk neutral individuals, the indifference condition for the marginal entrepreneur then becomes

$$\pi^d(\varphi^d) = (1 - U)\bar{w} = s. \tag{3.8$'$}$$

Together, Eqs. (3.23) and (3.8′) imply that (only) the lowest-paid manufacturing workers, employed by the marginal firm with productivity φ^d, are paid the same wage as workers in the service sector. Hence, all production workers employed by infra-marginal firms hold "good" jobs in the sense that they get wages in excess of the wage rate in the service sector.

In comparison to the full employment version of our model, the relative operating profits of more productive firms are lower with rent-sharing, since part of the advantage stemming from higher productivity is compensated by having to pay a higher wage rate. Formally, the elasticity of firm-level relative operating profits with respect to relative firm productivity (cf. Eq. (3.7)) is no longer given by $\xi \equiv \sigma - 1$, but by $\bar{\xi} \equiv (\sigma - 1)/[1 + \theta(\sigma - 1)]$, which is smaller than ξ if θ is strictly positive.[27] It then follows from Eq. (3.23) that the elasticity of the firm-level wage with respect to firm-level productivity is given by $\theta\bar{\xi}$, while the elasticity of firm-level employment with respect to firm-level productivity is given by $(1 - \theta)\bar{\xi}$.

constraint (3.23), and hence the firm has no incentive to manipulate the wage, but instead treats it parametrically at the equilibrium level $w(\varphi) = \hat{w}(\varphi)$.

[25]Guided by the findings of Katz and Summers (1989), we maintain the assumption that the wage in the perfectly competitive service sector is fully flexible, and hence it is only the occupation as production worker which carries an income risk in our model.

[26]Production workers would of course prefer to work for a firm that offers higher wages and, in the absence of unemployment compensation, those who do not have a job would clearly benefit from working for any positive wage rate. However, since due to contractual imperfections it is impossible to fix effort of workers ex ante, firms are not willing to accept underbidding by outsiders: once employed, the new workers would adopt the reference wage of insiders and thus reduce their effort when the wage paid by the firm falls short of the wage considered to be fair (see Fehr and Falk, 1999).

[27]In the limiting case $\theta = 0$, firm-level operating profits have zero weight in the determination of the fair wage, Eq. (3.23) simplifies to $\hat{w} = w$, and the model collapses to the full employment version.

All results derived in earlier parts of this chapter are robust with respect to our extension featuring an imperfectly competitive labour market for production workers. In particular, the two counteracting effects of offshoring on firm-level employment do not change qualitatively. Of course, there are quantitative effects, which can be best understood by considering the following mechanism that additionally arises due to firm-level rent sharing: For an offshoring firm, there is a feedback effect on firm-level marginal costs in the source country, since higher operating profits lead to higher firm-level wage rates via fair-wage constraint (3.23). This implies that the input ratio changes more strongly in favour of the imported routine task. As a consequence, the international relocation effect identified in Eq. (3.19) is now multiplied by the factor $\xi/\bar{\xi} > 1$, and hence more strongly negative than in the full-employment model. In addition, the functional relationships between χ and the two inequality measures in Section 3.2 on the one hand and between χ and welfare on the other hand are still given by Eqs. (3.20) to (3.22), with the mere difference that $\bar{\xi}$ replaces ξ and $\bar{\zeta} \equiv k/(k-\bar{\xi})$ replaces ζ.[28] Hence, the comparative static effects of offshoring on aggregate welfare and on income inequality among entrepreneurs as well as between entrepreneurs and workers change only quantitatively, but remain qualitatively the same in the model extension considered here.

In the model variant with an imperfectly competitive labour market there are two further aggregate variables that are worthwhile to look at: involuntary unemployment and wage inequality among employed production workers. In the presence of firm-level rent sharing, L is the mass of individuals looking for employment as production workers in the source country, while the mass of (employed) production workers is now given by $(1-U)L$. Neither entrepreneurs nor workers in the offshoring-service sector can be unemployed, and therefore the economy-wide unemployment rate in the source country is given by $u \equiv UL/N$. When looking at u/u^a, it is helpful to consider separately the effect of offshoring on the unemployment rate of production workers, measured by U/U^a, and the effect on the supply of production labour due to adjustments in the occupational choice, measured by L/L^a.[29] As shown in Appendix A.8 , the unemployment rate

[28] A detailed discussion on how firm-level rent-sharing alters the equations in Section 3.1 is deferred to Appendix A.7.

[29] The importance of occupational choice for understanding how a country's labour market absorbs the consequences of trade and offshoring has recently been pointed out by Liu and Trefler (2011) and Artuç and McLaren

of production workers is given by

$$U = \frac{\theta(\bar{\zeta}-1)+1-\Delta(\chi;\eta)}{\theta(\bar{\zeta}-1)+1}, \tag{3.24}$$

where $\Delta(\chi;\eta) \equiv \beta(\chi;\eta)/\alpha(\chi;\eta)$ and

$$\beta(\chi;\eta) \equiv 1+\chi^{\frac{k-(1-\theta)\bar{\xi}}{k}}\left[\eta\left(1+\chi^{\frac{\bar{\xi}}{k}}\right)^{(1-\theta)}-1\right], \quad \alpha(\chi;\eta) \equiv 1+\chi^{\frac{k-\bar{\xi}}{k}}\left[\eta\left(1+\chi^{\frac{\bar{\xi}}{k}}\right)-1\right]. \tag{3.25}$$

It is easily checked that $\Delta(0,\eta)=1$, and therefore U is lower in an equilibrium with offshoring than in autarky if $\Delta(\chi;\eta)>1$ and higher than in autarky if $\Delta(\chi;\eta)<1$. The effect of offshoring on L follows directly from Eq. (3.15), and as discussed in Subsection 3.2.1, the supply of production labour is smaller in an offshoring equilibrium than in autarky. By reducing L, this effect reduces aggregate unemployment u, ceteris paribus. Putting together these partial effects leads to

$$\frac{u}{u^a} = \Lambda(\chi;\eta), \quad \text{with} \quad \Lambda(\chi;\eta) \equiv \frac{\theta(\bar{\zeta}-1)+1-\Delta(\chi;\eta)}{\theta(\bar{\zeta}-1)}\frac{[1+\bar{\zeta}(\sigma-1)]\gamma}{1+\bar{\zeta}(\sigma-1)\gamma}, \tag{3.26}$$

where u^a can be computed from Eqs. (3.15) and (3.24). The first fraction of $\Lambda(\chi;\eta)$ is equal to U/U^a and the second fraction is equal to L/L^a. Unemployment rate u is lower with $\chi>0$ than with $\chi=0$ if $\Lambda(\chi;\eta)<1$, while the opposite is true if $\Lambda(\chi;\eta)>1$. We show the following result.

Proposition 3.3.1 *Unemployment in the source country decreases with the share of offshoring firms at low levels of χ. Under the sufficient condition*

$$\eta > \hat{\eta} \equiv \frac{2^\theta\theta\bar{\xi}}{2^\theta\theta\bar{\xi}+(2^\theta-1)(k\sigma-\bar{\xi})}$$

the effect is reversed as more firms offshore, and unemployment surpasses its autarky level if χ is sufficiently large.

Proof See Appendix A.9, where it is also shown that $\hat{\eta}<0.5$ if $k \geq 2$.[30]

(2012).

[30]Since empirical estimates for k are higher than two, it follows that, when focusing on the empirically relevant parameter domain, $\eta>0.5$ is sufficient for unemployment in the neighbourhood of $\chi=1$ being higher than under autarky.

The intuition for this result is straightforward. Since the labour supply effect works unambiguously in favour of a reduction in overall unemployment, cf. Eq. (3.15), all potentially harmful employment effects must work via an increase in the unemployment rate of production workers U. This effect is understood most easily by noting that the fair-wage constraint implies $w^d(\varphi^d) = \pi^d(\varphi^d)$, which together with the indifference condition for the marginal entrepreneur leads to

$$U = 1 - \frac{w^d(\varphi^d)}{\bar{w}}$$

in any equilibrium with $\chi < 1$. Whenever the average wage of employed production workers is higher than the wage paid by the marginal firm (which is the case whenever there is firm-level rent sharing) this is accompanied in equilibrium by a strictly positive level of unemployment.

Moreover, we see that if $\bar{w}/w^d(\varphi^d)$ changes, U has to change in the same direction, which has the following implication: For an increase in χ, starting from zero the international relocation effect in Eq. (3.19) dominates and offshoring displaces workers in high-productivity firms, which – due to the rent-sharing mechanism – earn high wages, thereby reducing the average wage relative to the wage paid by the marginal firm. This is only compatible with indifference between occupations if unemployment of production workers decreases as well. The effect of a marginal increase in offshoring on U is reversed at high levels of χ, since now the productivity effect in Eq. (3.19) is dominant, such that both newly offshoring *and* infra-marginal offshoring firms create additional high-wage jobs, pushing up the average wage relative to the wage paid by the marginal firm, which is only compatible with indifference between occupations if unemployment of production workers increases as well. Overall unemployment is then driven by two opposing effects: the supply of production workers decreases, but a larger share of them is without a job. If η is large, and hence the international relocation effect is small, the negative impact of offshoring on U dominates the decline in L at high levels of χ.

The ratio $\bar{w}/w^d(\varphi^d)$ provides one measure of income inequality among production workers, but not a very informative one, since it ignores information on individual wage rates by everybody but the workers in the marginal firm. Hence, in analogy to the measurement of entrepreneurial income inequality we now look at the Gini coefficient as a more sophisticated measure of wage dispersion. As formally shown in Appendix A.10, this Gini coefficient is given by

$$A_L(\chi) = \frac{\theta(\bar{\zeta}-1)}{2+\theta(\bar{\zeta}-1)} \left\{ 1 + \frac{2\left(1-\chi^{\frac{k-(1-\theta)\bar{\xi}}{k}}\right)[\alpha(\chi;\eta)-1]}{\alpha(\chi;\eta)\beta(\chi;\eta)\theta(\bar{\zeta}-1)} - \frac{2\left[1+\theta(\bar{\zeta}-1)\right]\left(1-\chi^{\frac{k-\bar{\xi}}{k}}\right)[\beta(\chi;\eta)-1]}{\alpha(\chi;\eta)\beta(\chi;\eta)\theta(\bar{\zeta}-1)} \right\}. \quad (3.27)$$

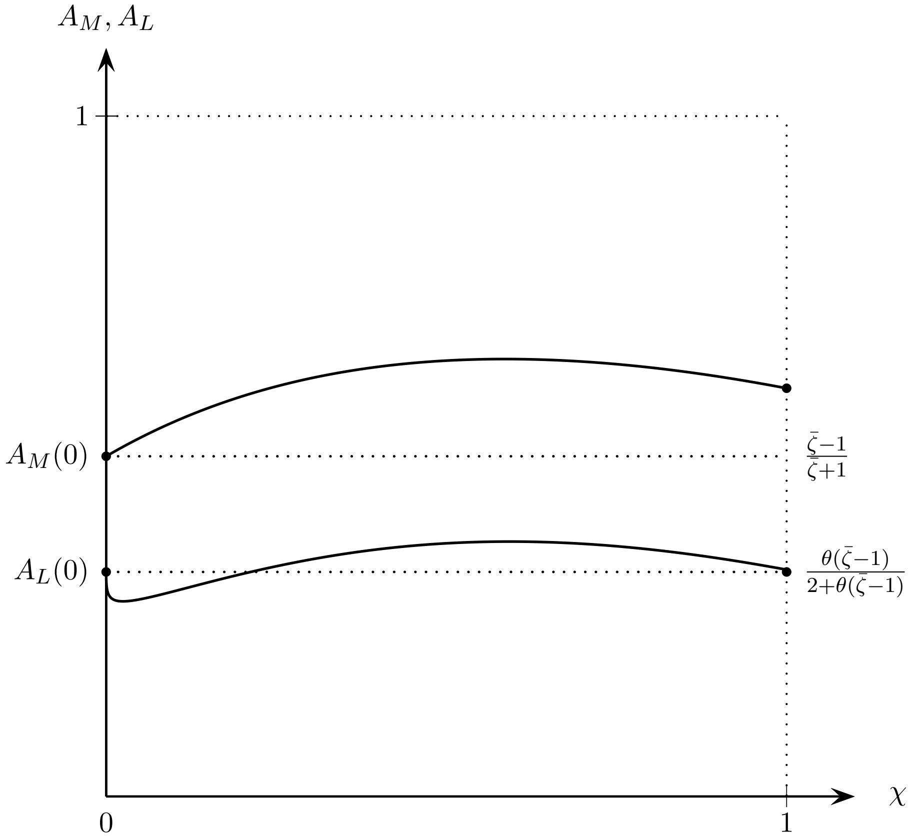
Inequality of wage income is the same in the polar cases where either no firms or all firms offshore: $A_L(0) = A_L(1) = \theta(\bar{\zeta}-1)/[2+\theta(\bar{\zeta}-1)]$.[31] We can furthermore show that A_L is lower than the autarky level at low levels of offshoring, and higher than the autarky level at high levels of offshoring. Figure 3.3 illustrates the resulting S-shape of the A_L locus, alongside the Gini-coefficient for entrepreneurial income A_M that we computed in Section 3.2.2, with the only modification that now $\bar{\zeta}$ replaces ζ.

The intuition is analogous to the one for the effect of offshoring on $\bar{w}/w^d(\varphi^d)$. In a situation where the offshoring strategy is only chosen by the most productive firms, the international relocation effect shifts good (high-wage) jobs abroad, and displaced workers have to accept less well paid jobs in- *and* outside the manufacturing sector. This effect is in accordance with results from Ebenstein, Harrison, McMillan, and Phillips (2013), who find for the US that workers who have to switch occupations as they are displaced from the manufacturing sector suffer discrete income losses of about 12 to 17 percent, and in our model it is responsible for the reduction of wage inequality at low levels of χ. The influence of the relocation effect is reversed at high levels of χ, since now the low-productivity firms shift low-wage jobs abroad, thereby contributing to an increase in wage inequality in the source country. There is also a firm-level wage effect due to the rent-sharing mechanism in our model: It increases wage dispersion at low levels of χ (wage-boosting increase in profits by high-wage firms) and reduces wage dispersion at high levels of χ (wage-boosting increase in profits by low-wage firms). The firm-level wage effect thereby influences wage inequality in the opposite direction to the international relocation effect, and it dominates the overall effect when many firms offshore.[32]

[31] An analogous result holds for the trade models of Egger and Kreickemeier (2009, 2012) and Helpman, Itskhoki, and Redding (2010), where wage inequality is the same in the cases of autarky and exporting by all firms.

[32] As we show in Appendix A.12, the Gini coefficient for the income distribution within the broadly defined group

Figure 3.3: *Gini coefficients for entrepreneurial income and wage income*



The following proposition summarises the main insights regarding the distributional effects of offshoring within the group of (employed) production workers.

Proposition 3.3.2 *The impact of offshoring on the dispersion of wage income, measured by the Gini coefficient, is non-monotonic. Wage income inequality falls relative to the benchmark without offshoring if χ is small, while it rises relative to this benchmark if χ is sufficiently large.*

Proof Analysis in the text and formal discussion in Appendix A.11.

of *all* production workers, including those who are unemployed, is given by $A_U(\chi) = [1 - U(\chi)] A_L(\chi) + U(\chi) \geq A_L(\chi)$. Since $U(\chi)$ is smaller than $U(0)$ at low levels of χ, while the reverse is true at high levels of χ, the non-monotonic effect of χ on $A_L(\chi)$ is reinforced. The only difference in the behaviour of both indices is that $A_U(1) > A_U(0)$ while $A_L(1) = A_L(0)$, which results from the fact that $U(1) > U(0)$.

3.4 Economy-wide inequality

So far, our focus was on inequality within and between various subgroups of the population. We now analyse the impact of offshoring on economy-wide inequality. For computing a comprehensive measure of economy-wide income inequality, we have to solve the problem that distributions of profit and labour income overlap if $\theta > 0$. Due to this overlap, we cannot simply calculate Gini coefficients for ranking the economy-wide income distributions with and without offshoring, but instead look at the Theil index as an alternative measure of income inequality. In discrete notation, the Theil index for the income distribution in a group of agents with population size n can be computed according to

$$T = \frac{1}{n} \sum_{i=1}^{n} \frac{y_i}{\bar{y}} \ln \left(\frac{y_i}{\bar{y}} \right), \tag{3.28}$$

where y_i is income of agent i, while $\bar{y}$ is the average income. If income is equally distributed, the Theil index has a value of zero. The index increases with inequality and reaches a maximum value of $\ln n$ if all the income is fully concentrated on one person. This implies that the range of the Theil index depends on population size. One of the main advantages of the Theil index as compared to other measures of inequality is its decomposability. For instance, if there are m subgroups of population, Theil index T can be decomposed according to

$$T = \sum_{j=1}^{m} \frac{n_j \bar{y}_j}{n \bar{y}} T_j + \sum_{j=1}^{m} \frac{n_j \bar{y}_j}{n \bar{y}} \ln \left(\frac{\bar{y}_j}{\bar{y}} \right), \tag{3.29}$$

where $\sum_{j=1}^{m} n_j = n$ and T_j refers to the Theil index of income group j, which can be computed in analogy to Eq. (3.28). The Theil index can thus be written as a weighted average of inequality within subgroups, plus inequality between these subgroups (cf. Shorrocks, 1980). This makes it particularly useful for our purpose.

In our model, we can distinguish between self-employed agents (entrepreneurs plus offshoring-service agents) and all production workers (employed and unemployed) as the two main income groups. Denoting the Theil indices for these specific subgroups by T_S and T_U, respectively, the

Theil index for the economy-wide income distribution can be written as[33]

$$T = a_S \left(T_S + \ln \bar{\zeta} \right) + a_U T_U + \ln \left(\frac{a_S}{\bar{\zeta}} + a_U \right), \tag{3.30}$$

with

$$a_S \equiv \frac{1-\rho}{\rho\gamma + 1 - \rho} \qquad \text{and} \qquad a_U \equiv \frac{\rho\gamma}{\rho\gamma + 1 - \rho} \tag{3.31}$$

being the income shares of the two population subgroups. To understand how offshoring influences Theil index T, we first look at the benchmark scenario without firm-level rent-sharing, i.e. $\theta = 0$. In this case, all firms pay the same wage and all production workers find a job, implying that Theil index T_U falls to zero. Eq. (3.30) therefore simplifies to

$$T = \frac{1-\rho}{\rho\gamma + 1 - \rho} \left(T_S + \ln \zeta \right) + \ln \left[\frac{\zeta\rho\gamma + 1 - \rho}{\zeta(\rho\gamma + 1 - \rho)} \right], \tag{3.32}$$

where $\bar{\zeta}$ has been replaced by ζ due to $\theta = 0$. To see how offshoring affects economy-wide inequality, it is crucial to understand how it influences the distribution of income within the subgroup of self-employed agents. From the analysis in Section 3.2.2 we already know that offshoring raises inequality within the group of self-employed agents according to the Gini criterion. However, this is not sufficient for an increase in Theil index T_S. Unlike the Gini coefficient, the Theil index does not rely on the Lorenz curve. But the two indices share one important property: both of them respect mean-preserving second-order stochastic dominance, which is equivalent to Lorenz dominance. We can therefore conclude that the Gini coefficient and the Theil index rank two distributions equivalently, if one of them Lorenz dominates the other one. One can show that the distribution of income among self-employed agents under autarky Lorenz dominates the respective distribution in an offshoring equilibrium for arbitrary values of $\chi \in (0, 1)$.[34] This implies $T_S > T_S^a$ (where superscript a refers to autarky).

Accounting for $T_S > T_S^a$, it follows from Eq. (3.32) that $T - T^a > \Delta_T(\chi)$, with[35]

$$\Delta_T(\chi) \equiv \frac{\rho(1-\rho)(1-\gamma)(\zeta - 1)}{\rho\gamma + 1 - \rho} + \ln \left[\frac{\zeta\rho\gamma + 1 - \rho}{\zeta(\rho\gamma + 1 - \rho)} \right] - \ln \left[\frac{1 + \rho(\zeta - 1)}{\zeta} \right].$$

[33] See Appendix A.13 for derivation details.

[34] Showing Lorenz dominance in this case is tedious, and therefore we have delegated formal details of this analysis to Appendix A.14.

[35] Thereby, $T_S^a = (\zeta - 1)^{-1} \int_1^\infty x^{-k/\xi} \left[\ln x - \ln \zeta \right] dx = \zeta - 1 - \ln \zeta$ has been considered.

Economy-wide inequality is higher with offshoring than under autarky if $\Delta_T(\chi) > 0$ holds for $\chi > 0$. This is the case, because

$$\frac{d\Delta_T(\chi)}{d\chi} = -\frac{\rho^2(1-\rho)\gamma(\zeta-1)^2}{[\rho\gamma+(1-\rho)]^2(\zeta\rho\gamma+1-\rho)}\frac{d\gamma(\chi;\eta)}{d\chi} > 0$$

and $\Delta_T(0) = 0$. We can therefore conclude that an increase of χ from zero to any positive level increases Theil index T, and hence renders the economy-wide distribution of income less equal.

Things are more complicated if rent sharing gives rise to firm-specific wages and involuntary unemployment, because in this case changes in Theil index T additionally account for adjustments in the distribution of income within the group of production workers, as captured by T_U. Since we know from the analysis in the previous section that offshoring may increase or decrease income inequality among production workers, it is *a priori* not clear, whether offshoring in our model renders the economy-wide distribution of income more or less even than under autarky.[36] We address this question as part of the numerical exercise conducted in the next section.

3.5 A quantitative exercise

In this section, we conduct a numerical exercise using the model variant from Section 3.3 and parameter estimates from the empirical trade literature. The purpose of this exercise is twofold. On the one hand, our aim is to illustrate the non-monotonic effect of offshoring on inequality, welfare, and unemployment. On the other hand, we want to shed additional light on the consequences of offshoring for the economy-wide distribution of income under firm-level rent-sharing, for which the analytical results are not clear. Since our model – even with its extension including firm-level rent sharing – is highly stylized, the quantitative effects should be viewed as illustrative.

A first set of parameters is taken from Egger, Egger, and Kreickemeier (2013), who structurally estimate key parameters of a trade model along the lines of Egger and Kreickemeier (2012), which is in may respects similar to the theoretical framework underlying our analysis, but does not account for offshoring. Employing information from the Amadeus data-set, Egger,

[36]In Appendix A.15, we show that a movement from autarky to high levels of offshoring increases economy-wide inequality if θ is sufficiently small.

Egger, and Kreickemeier (2013) report the following parameter estimates for the average country in their data-set, which covers five European economies: $\theta = 0.102$, $\sigma = 6.698$, $k = 4.306$. While, to the best of our knowledge, there are no other directly comparable estimates for the rent-sharing parameter available, the estimate of σ lies in the range of parameter estimates reported by Broda and Weinstein (2006) and is well in line with the parameter value considered by Arkolakis (2010) in his calibration exercise. The parameter estimate of k is higher than the estimate of about 2 reported by Corcos, Gatto, Mion, and Ottaviano (2012). However, it is consistent with findings by Arkolakis and Muendler (2010) and – together with the estimates for θ and σ – guarantees that the parameter constraint $k > \bar{\xi}$ is fulfilled.

It is challenging to come up with a theory-consistent measure of η. We take guidance from the findings by Blinder (2009) and Blinder and Krueger (2013) that about a quarter of jobs in US manufacturing can be classified as offshorable.[37] In our model, all jobs done by individuals employed in routine tasks can in principle be offshored, and the economy-wide cost share of these jobs is $1 - \eta$. Under autarky, all workers within each firm are paid the same wage, and therefore in this situation $1 - \eta$ is also the fraction of jobs that can be offshored. We therefore set $\eta = 0.75$. Of course, estimates on the actual extent of offshoring are much smaller than the numbers reported by Blinder (2009) and Blinder and Krueger (2013). For the US economy, Forrester Research predicted in 2002 a loss of 3.3 million jobs due to offshoring by 2015, which is less than 2.5 percent of the workforce. Bringing the quantitative results from our numerical exercise in accordance with such estimates therefore requires that the fraction of offshoring firms is sufficiently small.

Based on these parameter estimates, we can quantify the effects of offshoring. For this purpose, we compute how a given exposure to offshoring alters our variables of interest relative to a benchmark without offshoring. Thereby, we first look on changes of intra- and inter-group inequality and determine the relative importance of these changes for the adjustments in economy-wide inequality. The results from this exercise are summarised in Table 3.1.

[37] Based on the taxonomy of Blinder (2009) and Blinder and Krueger (2013) researchers have provided estimates on the share of offshorable tasks also for other industrialised countries. For Germany, the share of jobs that can be classified offshorable amounts to 42 percent and is thus significantly higher than for the US (see Laaser and Schrader, 2009).

Table 3.1: *Impact of offshoring on different measures of inequality*

	Change of					
χ	A_M in pct.	A_L in pct.	$\bar{\omega}$ in pct.	T_S in pct.	T_U in pct.	T in pct.
0.001	0.033	−8.685	0.084	0.167	−9.818	−0.174
0.01	0.322	−6.910	0.837	1.174	−9.161	2.589
0.10	2.860	−0.488	8.369	6.485	−1.291	11.642
0.25	5.902	2.270	20.922	10.735	3.953	18.499
0.50	8.626	2.362	41.844	13.762	9.351	23.967
0.75	9.395	1.224	62.766	14.820	12.062	26.488
0.90	9.211	0.477	75.319	15.025	13.035	27.319

Notes: All reported figures refer to percentage changes relative to autarky.

Columns 2 to 4 quantify the impact of offshoring on the inequality measures discussed in Sections 3.2 and 3.3. Evaluated at our parameter estimates, offshoring has only a moderate effect on intra-group inequality among (employed) production workers and among entrepreneurs, whereas its impact on inter-group inequality between entrepreneurs and production workers can be sizable. Columns 5 to 7 summarise the quantitative effects of offshoring on the distribution of income within the two main income groups – self-employed agents (entrepreneurs plus offshoring service agents) and all production workers (employed and unemployed) – as well as for the whole economy, relying on Theil indices. The qualitative effects of offshoring on income inequality within the now more broadly defined income groups are the same as those reported in Columns 2 and 3, but the quantitative effects seem to be more pronounced. The quantitative differences regarding the effects of offshoring on intra-group inequality can be explained by different definitions of income groups and by the fact that the Gini coefficient is confined to the unit interval, while this is not the case for the Theil index. Our numerical results point to a considerable increase in economy-wide income inequality at higher levels of χ. We can also see that, evaluated at our parameter estimates, offshoring lowers economy-wide income inequality if χ is sufficiently close to zero.

We now turn to the impact of offshoring on welfare and unemployment, which we summarise in Table 3.2. As outlined in Section 3.2.3, the impact of offshoring on source-country welfare crucially depends on the value of ε. To illustrate this, we run separate numerical experiments for the two polar cases highlighted in Section 3.2.3: a production technology without external increasing returns to scale ($\varepsilon = 0$) and a textbook CES production technology ($\varepsilon = 1$). The results for these two exercises are reported in Columns 2 and 3. Thereby, Column 2 confirms our analytical finding that in the absence of external increasing returns to scale source country income I declines relative to autarky at low levels of χ. However, the welfare loss is small compared to the potential welfare gains at high levels of χ. With a textbook CES production technology, external increasing returns to scale generate additional welfare gains from firm entry, and these gains are sufficiently strong to dominate welfare losses from unfavourable labour reallocations at low levels of χ. At higher levels of χ offshoring leads to firm exit and in this case the external increasing returns to scale viewed on their own lead to a welfare loss. However, this loss is not strong enough to dominate the positive welfare implications of the now more favourable resource allocation and relying on the textbook production technology offshoring is therefore a success story for the source country, irrespective of the level of χ.

Table 3.2: *Impact of offshoring on welfare and unemployment*

	Change of			
	I in pct.			u in ppt.
χ	$\varepsilon = 0$	$\varepsilon = 1$	$\varepsilon = 0.56$	
0.001	−0.800	0.624	−0.005	−2.554
0.01	−0.827	1.164	0.283	−2.654
0.10	2.067	3.796	3.032	−0.944
0.25	7.202	7.289	7.251	0.898
0.50	15.222	12.235	13.540	2.560
0.75	22.564	16.492	19.125	3.502
0.90	26.692	18.804	22.212	3.839

Notes: Welfare effects refer to percentage changes relative to autarky, whereas unemployment effects refer to changes in percentage points.

The results for the two cases $\varepsilon = 0$ and $\varepsilon = 1$ define a corridor in which the welfare effects of offshoring can lie in our model. We also provide results using $\varepsilon = 0.56$, which is the empirical estimate of Ardelean (2011). The insights from this exercise are summarized in Column 4, and we see that in this case there are small welfare losses from offshoring for the source country if $\chi = 0.001$. The last column of Table 3.2 confirms our analytical finding from Section 3.3 that offshoring lowers aggregate unemployment at low levels of χ, whereas it exacerbates the unemployment problem in the source country at high levels of χ. In general, the quantitative effect of offshoring on economy-wide unemployment is fairly small, when evaluating the model at our parameter estimates.

To complete our discussion on the quantitative effects of offshoring, we finally look more specifically on the consequences of the observed exposure to offshoring. This requires empirical information upon the share of firms that engage in offshoring, which is reported for Germany by Moser, Urban, and Weder di Mauro (2009). Using a large sample of 8,466 German plants from the IAB Establishment Panel, they find that the share of offshoring firms is 14.9 percent. This share is somewhat lower than the share of offshoring firms reported by The Economist (2004) from a small survey of 150 British firms, while it is significantly higher than the share of firms conducting international outsourcing and/or FDI in Japan as reported by Tomiura (2007). Evaluated at $\chi = 0.149$, our model predicts that offshoring has increased inequality within the group of entrepreneurs by 4.0 percent and inequality within the group of production workers by 0.9 percent, when relying on the metric of Gini coefficients. Looking at the relative income of entrepreneurs and workers, offshoring has augmented the preexisting gap in Germany by 12.5 percent. Also economy-wide income inequality has widened considerably due to offshoring, with the respective Theil index being 14.4 percent higher under the observed exposure to offshoring than under autarky. With respect to its welfare consequences, our model predicts a moderate increase for Germany, ranging between 3.1 (for $\varepsilon = 0$) and 5.0 (for $\varepsilon = 1$) percent. Using Ardelean's estimate of $\varepsilon = 0.56$, the welfare increase amounts to 4.5 percent. In contrast to the widespread perception of large negative employment effects, our model predicts that offshoring has lowered unemployment in Germany by 0.2 percentage points.[38]

[38]Since empirical evidence for Germany suggests setting $\eta = 0.58$ instead of $\eta = 0.75$, we have repeated our numerical exercise from this paragraph for $\eta = 0.58$. This change in the value of η does not affect the predicted consequences of observed offshoring for A_M and $\bar{\omega}$, and it has only a small quantitative effect on the

3.6 Summary

In this chapter we have developed an analytically tractable general equilibrium framework for analysing offshoring to low-wage countries. It is a key feature of our framework that firms differ from each other in terms of their productivity. As a consequence, the costly option to offshore routine tasks to the low-wage country, while available to all firms, is chosen only by a subset of them in equilibrium. The effects that offshoring has on welfare and the income distribution depends on the share of firms that offshore tasks in equilibrium, and we are therefore able to show that considering firm heterogeneity adds a relevant dimension to the established offshoring literature that has mainly focussed on the heterogeneity of tasks.

Offshoring is attractive for firms because it leads to lower marginal production costs, and this implies an expansion of employment in non-routine tasks at home. However, offshoring at the same time destroys domestic jobs in which workers perform routine tasks. The relative strength of these two opposing effects on domestic firm-level employment depends on the costs of offshoring. If these costs are high, offshoring is only attractive for a relatively small fraction of high-productivity firms, because its potential for lowering marginal production costs is small. As a consequence, the destruction of domestic routine-task jobs dominates the establishment of new jobs in which workers perform non-routine-tasks, and offshoring hence lowers domestic firm-level employment. Workers losing their jobs in offshoring firms find employment in less productive activities, including jobs in low-productivity firms newly entering the domestic market. Unlike trade in final goods, which in canonical models with heterogeneous producers triggers a reallocation of domestic workers from low- to high-productivity firms, offshoring therefore causes a shift of domestic employment from high- to low-productivity firms.

The reallocation of workers from low- to high-productivity firms constitutes a detrimental

predicted consequences for I. At the same time, using the lower value for η leads to larger quantitative effects of observed offshoring on economy-wide inequality and aggregate unemployment in Germany. With $\eta = 0.58$ our model predicts that offshoring has increased economy-wide inequality by 21.73 percent and has lowered aggregate unemployment by 2.46 percentage points. Finally, the reduction of η changes the predicted consequences of offshoring for A_L in a qualitative way. According to our model the observed exposure to offshoring has lowered intra-group inequality among production workers in Germany by 2.19 percent, when considering $\eta = 0.58$ instead of $\eta = 0.75$.

welfare effect, which can dominate traditional sources of welfare gain, and therefore render the source country worse off with offshoring than in autarky. The situation is more favourable at lower costs of offshoring, because in this case offshoring becomes attractive for a broad range of producers and leads to a reallocation of workers towards high-productivity firms. As a consequence, source country welfare unambiguously increases relative to autarky if the costs of offshoring are sufficiently small.

Income inequality between entrepreneurs and workers increases unambiguously with the share of offshoring firms. However, the effect on income inequality among entrepreneurs is non-monotonic: income inequality within this group increases if only a few firms shift the production of routine tasks abroad, and it decreases (while always staying above the autarky level) if offshoring becomes common practice among high- and low-productivity firms. Both of these effects contribute to the emergence of a new class of entrepreneurial superstars, who gain disproportionately from the global expansion of their firms under offshoring.

An extended version of our model with firm-level rent sharing, which preserves all results derived in the benchmark model with perfectly competitive labour markets, allows us to address the public concern that offshoring destroys domestic jobs and exacerbates the problem of unemployment. Our analysis shows that it is important to distinguish between what happens at the level of offshoring firms (firm-level effect) and what happens in the aggregate, after taking into account general equilibrium effects. We find that firm-level employment of production workers and aggregate employment tend do move in opposite directions: aggregate employment increases unambiguously at low levels of offshoring, where the negative firm-level effects on source country employment are largest. The reverse is true at high levels of offshoring: firm-level employment of production workers goes up, while aggregate employment falls.

The model extension with rent sharing also provides a richer picture of the distributional effects of offshoring, by additionally allowing for wage inequality of ex ante identical production workers. To understand its distributional consequences for production workers, it is noteworthy that offshoring constitutes a threat to the incomes of workers employed in both *good* (high-wage) and *bad* (low-wage) jobs. The former fear the relocation of their jobs abroad at early stages of offshoring, leaving them with alternatives that invariably yield lower incomes. The latter face a potential shut-down of their firms when production shifting becomes common practice among high- and low-productivity employers at later stages of offshoring, and some of those losing their job join the ranks of the unemployed. An immediate consequence of these firm-level employment effects is that offshoring reduces wage inequality initially, but widens it if a sufficiently large share of firms shifts the production of routine tasks abroad. A non-monotonicity also materializes with respect to the effect of offshoring on economy-wide inequality. Relying on the Theil index, we show that economy-wide inequality decreases if only a few high-productivity firms make use of offshoring, whereas it increases if offshoring also becomes common practice among firms with lower productivity levels.

Our analysis highlights the relevance of the extensive margin of offshoring for understanding how relocating routine tasks to low-wage countries affects economy-wide variables, such as income inequality, welfare, and unemployment. Firms in our model react differently to the offshoring opportunity, and we show that their asymmetric response has important general equilibrium effects. We hope that these insights together with the tractability of our framework can provide guidance to the rapidly growing empirical literature on offshoring using firm-level data, and that it will also be a useful point of departure for further theoretical work.

Chapter 4

Offshoring and individual skill upgrading

It is a common feature of advanced economies that their workforces are increasingly engaged in the performance of more complex production tasks. Along with this changing structure of skill requirements, individuals constantly retrain and update their capabilities. According to Eurofound's European Working Conditions Survey 2010 (cf. Eurofound, 2012), industry-wide on-the-job training rates in Germany have increased from on average 28.4% in 2005 to about 40% in 2010. At the same time, more and more firms find it optimal to restructure their production processes by relocating the performance of offshorable tasks to low-wage countries abroad. Data from the OECD STAN bilateral trade data base show that the output share of intermediate imports from non-OECD countries in German manufacturing has increased by a remarkable 62% over the same time span. In this chapter, we argue that both phenomena are linked. We offer a theory to explain the mechanism behind this link and an empirical analysis to show its significance and magnitude.

In general a positive link between offshoring and training should not come as a surprise since offshoring, which is associated with the relocation of tasks to low-wage countries abroad, in the end (at least temporary) displaces some workers from their jobs. As shown by Hummels, Munch, Skipper, and Xiang (2012), workers who are displaced because of offshoring have a particularly high probability to acquire vocational training during the subsequent period of transitional unemployment. We add to this literature, focusing instead on the impact that offshoring has on currently *employed* individuals and not only on those who directly loose their job through offshoring. This new focus is motivated by two facts: On the one hand, the number of workers, which are directly displaced from their job by offshoring, is dwarfed by the mass of individuals, which stay in their job.[1] On the other hand, it is well known from the theoretical trade literature that offshoring not only leads to direct job losses for workers whose tasks are shifted abroad, but also has a (positive) *productivity effect*, which may benefit *all* workers, as firms pass through productivity gains from offshoring to domestic workers in form of higher wages (Kohler, 2004b; Grossman and Rossi-Hansberg, 2008; Rodriguez-Clare, 2010). It is exactly this productivity effect which in our theoretical model creates incentives for on-the-job training by increasing the associated wage gain of workers beyond the cost of skill upgrading.

To structure our idea, we set up a small-open-economy model of offshoring in the spirit of Grossman and Rossi-Hansberg (2008), featuring two offshorable sets of tasks, which differ in their skill requirements. Unlike in standard trade models, where endowments are fixed, workers in our model may react to a given offshoring shock by selecting into costly on-the-job training, thereby gaining abilities that are needed to perform skill-intensive high-wage tasks. Since the productivity effect of offshoring (cf. Grossman and Rossi-Hansberg, 2008) proportionally scales up wages for both task sets, the gap between these wages increases as well, rendering on-the-job training more attractive for untrained workers, who select into skill upgrading as long as the (offshoring induced) gap in wages exceeds the associated cost of skill upgrading.

Focusing on this training indifference condition we translate our theoretical model into an empirically testable specification. We thereby – in line with our theoretical results – expect that offshoring leads to more observed on-the-job training at the individual level – a relationship that we can estimate within a standard Probit framework. Our offshoring variable thereby relates to sectoral imports of intermediate products, which are a widely used measure to proxy for industry-level offshoring in the empirical trade literature (cf. Feenstra and Hanson, 1996a, 1999; Geishecker and Görg, 2008; Baumgarten, Geishecker, and Görg, 2013). Using the industry-level

[1] For example, in the sample of Hummels, Jørgensen, Munch, and Xiang (2013), only 9% of all workers observed from 1998 to 2006 loose their job through mass-layoff events. Out of those layoffs, again only 10% can be associated with increased offshoring by the respective employers.

variation in our offshoring measure to identify the impact on individual skill upgrading has the clear advantage that offshoring growth can be seen as *exogenous* to single workers, whose individual training decisions should not feed back into industry-level offshoring growth rates. This approach embeds our analysis into a recent and growing literature, which uses industry-level variation in globalization measures to identify effects that arise at the individual level (cf. Geishecker and Görg, 2008; Baumgarten, Geishecker, and Görg, 2013; Ebenstein, Harrison, McMillan, and Phillips, 2013). Data on individual skill upgrading decisions come from the "BIBB/BAuA Employment Survey 2005/06", which holds detailed information on individual participation in on-the-job training measures. Crucially, due to the high resolution of our data we can take into account a wide range of control variables, which in the empirical training literature (cf. Arulampalam, Booth, and Bryan, 2004; Bassanini, Booth, Brunello, De Paola, and Leuven, 2007) already have been identified as major determinants of individual skill-upgrading decisions. Of particular interest for our application is thereby the possibility to observe the introduction of technological innovations directly at the workplace, which gives us the opportunity to separate the effect of offshoring from the one of sectoral biased technological change (cf. Feenstra, 2010).

Our findings offer clear support for the mechanism laid out in our theoretical model. Offshoring growth has a positive and significant impact on the individual on-the-job training propensity of workers employed in German manufacturing between 2004 and 2006. This link holds for a number of specifications and is robust to the inclusion of various controls at the individual, firm, and industry level. After taking account of, among other things, technological change, business cycle effects, and firm-size differences, a one standard deviation higher offshoring growth at the industry level over the period 2004 to 2006 is related to an increase in the propensity to observe individual on-the-job training by between 3 to 7 percentage points.

This chapter connects two strands of the empirical literature, which so far mostly have been analysed in complete isolation. On the one hand we contribute to a literature that seeks to identify the determinants of individual on-the-job training decisions (see Bassanini, Booth, Brunello, De Paola, and Leuven (2007) for an overview). On the other hand, we also add to the empirical trade literature, which focuses on the implications that offshoring has for domestic labour markets (see Baumgarten, Geishecker, and Görg (2013); Ebenstein, Harrison, McMillan, and Phillips (2013) for recent examples). The first strand of the literature usually focuses on a combination of product and/or labor market based explanations to explain individual on-the-job training decisions in a closed-economy setting, thereby ignoring the impact that globalization may have on individual training decisions.[2] The empirical trade literature, on the contrary, mainly is concerned with the impact that offshoring has on skill upgrading in the aggregate. As a central result, several studies have shown that increased offshoring is associated with a rise in the share of high-skilled employment in total employment (cf. Crinò, 2008; Feenstra, 2010; Davies and Desbordes, 2012). Individual skill levels thereby usually are considered as fixed such that all skill upgrading takes place at the *extensive* margin between rather than at the *intensive* margin within workplaces. As a notable exception Hummels, Munch, Skipper, and Xiang (2012) show that workers, who are directly displaced from their job through offshoring are more likely to select into training measures before taking up a new job. We complement this research by focusing on the vast majority of workers staying in their jobs that, hence, are indirectly affected through the general-equilibrium effects of offshoring – effects to which they respond by increased on-the-job training.

The chapter is structured as follows. In the next section, we develop our theoretical model and derive as main prediction that offshoring growth leads to more individual skill upgrading. Subsequently, we look for the proposed link in the data and present an empirical analysis, which includes a description of the econometric set-up, the data used, the results obtained and a discussion on the timing and the robustness of the link between offshoring and on-the-job training. A final section concludes the chapter.

[2] Arulampalam, Booth, and Bryan (2004) and Bassanini, Booth, Brunello, De Paola, and Leuven (2007) control for a comprehensive range of individual-level indicators to explain the selection of workers into on-the-job training. Méndez and Sepúlveda (2012) point to the influence of the business cycle on skill upgrading and discuss carefully the different training-types and their respective business cycle properties. Additionally, Görlitz and Stiebale (2011) look at industry-level competition as a driver of on-the-job training decisions.

4.1 A simple model of offshoring and on-the-job training

The goal of this section is to describe an intuitive mechanism, which links offshoring and on-the-job training. To this end, we employ a simplified version of the Grossman and Rossi-Hansberg (2008) model of trade in tasks, focusing on a single industry, which produces a homogeneous, constant returns to scale output Y at a given world market price normalised to $p \overset{!}{=} 1$. The production of final output Y requires the performance of two task sets, $\widetilde{S}$ and $\widetilde{N}$, such that our production technology may be summarized by $Y = F(\widetilde{S}, \widetilde{N})$, with $\widetilde{S}$ and $\widetilde{N}$ replacing the usual inputs in the otherwise standard neoclassical production function $F(\cdot)$. The task sets, $\widetilde{S}$ and $\widetilde{N}$, differ in their skill requirements: While workers performing the $\widetilde{S}$-set must have some task-specific skills, no such skills are needed to perform tasks from the $\widetilde{N}$-set. For simplicity, we furthermore assume that both tasks sets consist of only two tasks: A non-offshorable task, S or N, and an offshorable task, S^* or N^*, which are combined according to technologies, $\widetilde{S} = \widetilde{S}(S, S^*)$ and $\widetilde{N} = \widetilde{N}(N, N^*)$.

The offshorable tasks, S^* or N^*, will be performed abroad, if the cost of doing so are sufficiently low, i.e. if $w_S \geq \tau_S w_S^*$ and $w_N \geq \tau_N w_N^*$, with $\tau_S, \tau_N \geq 1$ denoting the usual iceberg-type offshoring cost and w_S^* and w_N^* being the (constant) unit cost of performing the tasks S^* and N^* at a low-cost location abroad. The unit-costs for the task sets, $\widetilde{S}$ and $\widetilde{N}$, may then be written as $\omega_S(w_S, \tau_S w_S^*) = \Omega_S w_S$ and $\omega_N(w_N, \tau_N w_N^*) = \Omega_N w_N$, where $\Omega_S \equiv \omega_S(w_S, \tau_S w_S^*)/w_S \leq 1$ and $\Omega_N \equiv \omega_N(w_N, \tau_N w_N^*)/w_N \leq 1$ are defined as the cost savings factors from relocating tasks S^* or N^* abroad (cf. Grossman and Rossi-Hansberg, 2008). Analogously, the unit cost for final output Y may be expressed as $c(\Omega_S w_S, \Omega_N w_N) = \gamma c(w_S, w_N)$, with $\gamma \equiv c(\Omega_S w_S, \Omega_N w_N)/c(w_S, w_N) \leq 1$ denoting the total cost savings factor from (partly) offshoring both inputs used in $Y = F(\widetilde{S}, \widetilde{N})$.

We assume a homogeneous workforce of size $\bar{L} > 0$. Workers can either exclusively perform tasks from the $\widetilde{S}$-set or from the $\widetilde{N}$-set, whereas, as outlined above, tasks from the $\widetilde{S}$-set require task-specific skills, while no such requirement exists for tasks from the $\widetilde{N}$-set. To acquire the skills needed for the performance of tasks from the $\widetilde{S}$-set, workers have to invest in costly on-the-job training. Training cost $\kappa > 0$ (paid in units of the *numÃ©raire*) are assumed to be constant and workers invest into on-the-job training as long as the wage gain $w_S - w_N$ associated with it exceeds the corresponding cost κ. Accordingly, we may write the net gain from on-the-job training as

$$u \equiv w_S - w_N - \kappa \geq 0, \tag{4.1}$$

keeping in mind that in equilibrium $u = 0$ must hold, leaving workers indifferent between both alternatives.

Equilibrium wages under autarky (denoted by superscript a) and with offshoring (denoted by superscript o) can now be found in the intersection point of the training indifference condition Eq. (4.1) and the zero profit condition $\gamma c\,(w_N, w_S) = 1$ (see figure 1 below). As outlined above, $\gamma \leq 1$ thereby represents the total cost savings factor from offshoring, being equal to one under autarky and smaller than one in an equilibrium with offshoring.

In order to derive testable predictions on how offshoring alters wages and thus the training decision in Eq. (4.1), we have to specify our simple model in more detail. We assume that Y follows from a Cobb Douglas production technology, such that $F(\widetilde{S}, \widetilde{N}) = \widetilde{S}^{\alpha}\widetilde{N}^{1-\alpha}$ with $\alpha \in (0, 1)$. It then can be shown that the total cost savings from offshoring $\gamma = \Omega_S^{\alpha}\Omega_N^{1-\alpha} \leq 1$ are a weighted geometric mean of the cost savings at the task level, $\Omega_S \leq 1$ and $\Omega_N \leq 1$, respectively. The technology, according to which tasks within each of the two task sets are bundled together, is the same as in Antràs and Helpman (2004) and Acemoglu and Autor (2011). Assuming $\widetilde{S}\,(S, S^*) = BS^{\theta}\,(S^*)^{1-\theta}$ as well as $\widetilde{N}\,(N, N^*) = BN^{\theta}\,(N^*)^{1-\theta}$, with $\theta \in (0, 1)$ measuring the cost share of non-offshorable tasks and $B \equiv 1/[\theta^{\theta}\,(1-\theta)^{1-\theta}] > 0$ being a positive constant, we can infer that the cost savings from offshoring at the task-level, $\Omega_S = (\tau_S w_S^*/w_S)^{1-\theta} \leq 1$ and $\Omega_N = (\tau_N w_N^*/w_N)^{1-\theta} \leq 1$, are proportional to the respective international wage differential (including the transport costs τ_S or τ_N, respectively). An offshoring firm's profit maximization problem may hence be written as

$$\pi = \max_{\widetilde{S}, \widetilde{N}} \;\; F(\widetilde{S}, \widetilde{N}) - \Omega_S w_S \widetilde{S} - \Omega_N w_N \widetilde{N},$$

from which the corresponding first order conditions can be derived as

$$w_S\,(\tilde{s}) = f'\,(\tilde{s})/\Omega_S, \tag{4.2}$$

$$w_N\,(\tilde{s}) = [f\,(\tilde{s}) - \tilde{s}f'\,(\tilde{s})]/\Omega_N, \tag{4.3}$$

with $f\,(\tilde{s}) \equiv F(\widetilde{S}, \widetilde{N})/\widetilde{N} = \tilde{s}^{\alpha}$ referring to our production function in intensive form notation and $\tilde{s} \equiv \widetilde{S}/\widetilde{N}$ measuring the *overall* skill intensity in the entire production process (including domestic tasks, S and N, as well as foreign tasks, S^* and N^*).

From Eqs. (4.2) and (4.3), two channels through which offshoring affects domestic wages can be identified. As in Grossman and Rossi-Hansberg (2008), cost savings from offshoring are handed through to domestic workers in form of higher wages, which due to the *productivity effect* of offshoring are scaled up by factors, $1/\Omega_S \geq 1$ and $1/\Omega_N \geq 1$, respectively. On the contrary, the *labor supply effect* of offshoring leads to disparate wage effects by driving a wedge between the overall skill intensity $\tilde{s} \equiv \widetilde{S}/\widetilde{N}$, which applies for the entire production process, and the domestic skill intensity $s \equiv S/N$, which only reflects the composition of the domestic workforce. To illustrate the labor supply effect, Shephard's Lemma can be applied to $\omega_S(w_S, \tau_S w_S^*)$ and $\omega_N(w_N, \tau_N w_N^*)$, resulting in:

$$\frac{\partial \omega_S \left(w_S, \tau_S w_S^*\right)}{\partial w_S} \equiv \frac{S}{\widetilde{S}} = \theta \Omega_S \quad \text{and} \quad \frac{\partial \omega_N \left(w_N, \tau_N w_N^*\right)}{\partial w_N} \equiv \frac{N}{\widetilde{N}} = \theta \Omega_N. \tag{4.4}$$

Dividing both expressions in Eq. (4.4) by each other reveals how the domestic skill intensity, $s \equiv S/N$, is altered by the labor supply effect of offshoring, such that

$$\tilde{s} = \frac{\Omega_N}{\Omega_S} s, \tag{4.5}$$

emerges as the overall skill intensity. Intuitively, in the autarky equilibrium (with $\Omega_S = \Omega_N = 1$) the overall skill intensity coincides with the domestic skill intensity, implying $\tilde{s} = s$. With offshoring, the overall skill intensity additionally depends on which factor is offshored more intensively, such that $\tilde{s} \gtrless s$ if $N/\widetilde{N} \gtrless S/\widetilde{S}$. Intuitively, the labor supply effect of offshoring thereby favors the input factor which is offshored less intensively. When replacing $\tilde{s}$ in Eqs. (4.2) and (4.3) by Eq. (4.5) to determine the overall effect that offshoring has on domestic wages, it turns out that the productivity effect of offshoring is dominant and causes a proportional increase in *both* wages, $w_S^o(\tilde{s}) = w_S^a(s)/\gamma$ and $w_N^o(\tilde{s}) = w_N^a(s)/\gamma$, by the same factor $1/\gamma \geq 1$ for a notionally unchanged domestic factor intensity s.

To see the impact on workers' training decision, we may now substitute both wage rates into the training indifference condition (4.1), which then can be rewritten as:

$$u = w_S(s) - w_N(s) - \kappa = \frac{\alpha s^{\alpha-1} - (1-\alpha) s^{\alpha}}{\gamma} - \kappa, \tag{4.1$'$}$$

with $\gamma = \Omega_S^\alpha \Omega_N^{1-\alpha} < 1$ implying $s^o > s^a$. Intuitively, if both wages are scaled up by an identical factor $1/\gamma > 1$ the same holds true for the gap $w_S - w_N$ between these wages. In the end, as more and more domestic workers optimally react on $u > 0$ by upgrading their individual skills, the domestic skill intensity rises from s^a to s^o such that equilibrium is restored.

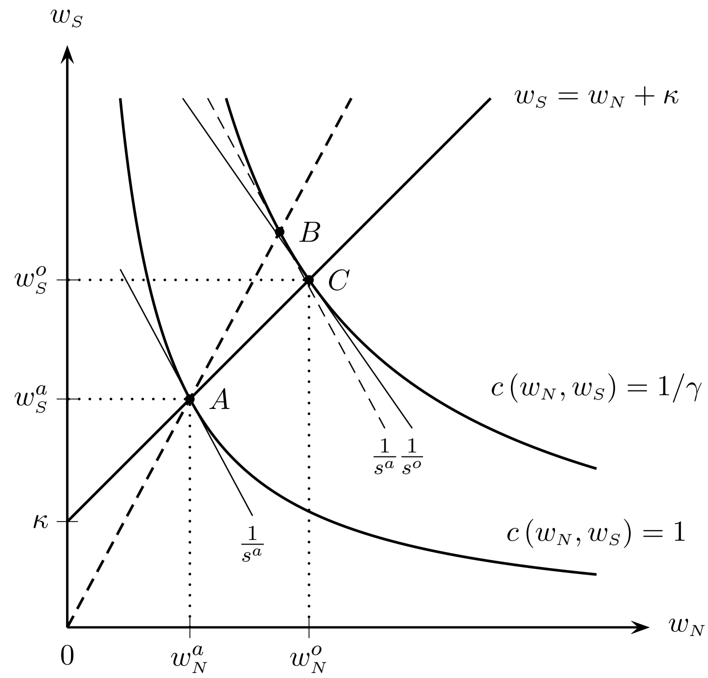


Figure 4.1: *On-the-job training with and without offshoring*

Figure 1 illustrates the effect of offshoring on on-the-job training. Starting out from the autarky equilibrium in A and holding the domestic skill intensity notionally fixed at $s = s^a$, offshoring causes a radial outward expansions of the unit-cost curve by factor $1/\gamma < 1$, which results in the hypothetical equilibrium B.[3] However, in point B we have $u > 0$, leaving domestic

[3]Fixing the domestic skill intensity at $s = s^a$ in this first step means that domestic workers are not allowed to switch tasks between the $\widetilde{N}$- and the $\widetilde{S}$-set. Of course this does not imply that workers are constrained in switching from offshorable N^*- or S^*-tasks to non-offshorable N- or S-tasks within the respective $\widetilde{N}$- or $\widetilde{S}$-set. Intuitively, the latter kind of task-arbitrage is a natural adjustment strategy to increased offshoring and a necessary condition for full-employment in our model.

workers with an incentive to invest in on-the-job training. As more and more workers decide in favor of on-the-job training, the domestic skill intensity increases from s^a to s^o until the new (offshoring) equilibrium C is reached. This result is at the heart of our analysis and we frame it in the following Proposition.

Proposition 4.1.1 *A decline in the cost of offshoring increases the share of tasks performed abroad, thereby leading to increased individual skill upgrading through on-the-job training.*

Proof Analysis in the text and formal discussion in Appendix A.17.

Summing up, offshoring positively impacts the individual decision in favor of on-the-job training. Interestingly, the training decision does not depend on the task content of offshoring. Even if only one task type is relocated abroad, $\Omega_S < 1$ or $\Omega_N < 1$ will be sufficient to induce $\gamma = \Omega_S^{\alpha}\Omega_N^{1-\alpha} < 1$ and, thus, more on-the-job training. Also note that offshoring not only affects the skill upgrading decision of those individuals which are directly hit by a (temporary) job loss though offshoring (cf. Hummels, Munch, Skipper, and Xiang, 2012). Rather it is the case that *all* individuals and in particular the vast majority of those who stay with their jobs are more likely to invest in individual skill upgrading as a response to given offshoring shock. Building upon these insights, we put Proposition 4.1.1 to the test by estimating the impact of increased industry-level offshoring on the individual on-the-job training decision displayed in Eq. (4.1).

4.2 The impact of offshoring on on-the-job training

The empirical part of this chapter is structured as follows: We lay out our empirical strategy in Subsection 4.2.1. Subsection 4.2.2 describes the data we use. The results of our empirical analysis then follow in Subsection 4.2.3. Finally, Subsections 4.2.4 and 4.2.5 discuss the timing of offshoring and skill upgrading and offer further robustness checks.

4.2.1 Empirical strategy

As a natural starting point to test Proposition 4.1.1, recall training indifference condition (4.1), which for individual $i = 1, ..., I$ employed in industry $j = 1, ..., J$ can be rewritten as

$$u_{ij} = w_{Sij} - w_{Nij} - \kappa_{ij}.$$

We know from Proposition 4.1.1 that any increase in offshoring (triggered by a decline in the offshoring costs τ_S or τ_N) widens the gap between w_{Sij} and w_{Nij}, thereby making on-the-job training more attractive for the individual worker. What we seek to identify in our empirical analysis is the realized on-the-job training in response to a given offshoring shock. We thus identify the adjustment mechanism described in our model above, according to which individuals engage in on-the-job training after an offshoring shock until a new equilibrium with $u_{ij} = 0$ and $s^o > s^a$ is reached. Unfortunately, an individual's net gain u_{ij} from on-the-job training is unobservable to us. Yet, we know that individual i selects into on-the-job training (indexed by $U_{ij} = 1$) if $u_{ij} > 0$ and does not do so (indexed by $U_{ij} = 0$) if $u_{ij} \leq 0$. We are thus able to portray the probability of on-the-job training as the outcome of an underlying latent variable model

$$Pr(U_{ij} = 1 \,|\, \cdot) = Pr(u_{ij} > 0 \,|\, \cdot), \tag{4.6}$$

conditioning on a vector $(\cdot)$ of observable covariates. Our main variable of interest is the growth rate of offshoring, $\widehat{O}_j$, in industry j, which, according to Proposition 4.1.1, should have a positive impact on the probability of on-the-job training in Eq. (4.6). We furthermore allow the individual training decision to depend on individual- and industry-specific characteristics, which we collect in vectors $\mathbf{Y}_i$ and $\mathbf{X}_j$, respectively. While the vectors $\mathbf{Y}_i$ and $\mathbf{X}_j$ will be specified in more detail below, we may for now interpret them as additional controls capturing such things as heterogeneity in the training cost κ_{ij}. Taken together, we can reformulate the training decision in Eq. (4.1) as:

$$u_{ij} = \beta_0 + \beta\widehat{O}_j + \mathbf{X}_j'\boldsymbol{\delta} + \mathbf{Y}_i'\boldsymbol{\eta} + \varepsilon_{ij}, \tag{1''}$$

with $\varepsilon_{ij} \sim N(0, 1)$ following a standard normal distribution with zero mean and variance one. We can then estimate the probability of on-the-job training $Pr(U_{ij} = 1 \,|\, \cdot)$ in Eq. (4.6) by a

Probit model based on the following empirical specification:

$$Pr\,(U_{ij}=1\,|\,\cdot\,) = Pr\,(u_{ij}>0\,|\,\cdot\,) = Pr(\beta_0 + \beta\widehat{O}_j + \mathbf{X}_j'\boldsymbol{\delta} + \mathbf{Y}_i'\boldsymbol{\eta} > \varepsilon_{ij}\,|\,\cdot\,). \qquad (4.6')$$

In line with Proposition 4.1.1, we expect a positive effect of offshoring growth $\widehat{O}_j$ on the probability of observing individual on-the-job training, i.e. $\beta > 0$. The identification of this relationship in our empirical model (4.6′) comes from varying offshoring growth rates across industries in which individuals are employed. This has the clear advantage that offshoring growth, which is measured at the industry level j, can be seen as exogeneous to worker i, whose individual training decision should not feed back into sector level offshoring growth. Consequently, we do not expect reverse causality to play a major role as potential source of endogeneity in our setting. This approach embeds our analysis into a recent and growing literature which uses industry level variation in globalization measures to identify individual level effects (Geishecker and Görg, 2008; Ebenstein, Harrison, McMillan, and Phillips, 2013; Baumgarten, Geishecker, and Görg, 2013). To limit the problem of omitted variables as another main reason for potentially biased estimates, we rely on a rich set of individual- and industry-specific covariates (summarized in $\mathbf{Y}_i$ and $\mathbf{X}_j$), which we introduce in Section 4.2.2 before discussing their role against the background of our empirical results in Section 4.2.3.

4.2.2 Data and definition of variables

Information on individual skill upgrading is taken from the "BIBB/BAuA Employment Survey 2005/06", which contains information on a wide set of workplace related variables for a representative sample of 20.000 individuals that participated between October 2005 and March 2006.[4] We use the latest available wave of what has become established as a reliable and detailed source for information related to on-the-job training (Acemoglu and Pischke, 1998; Dustmann and Schönberg, 2012). Our main dependent variable is the training incidence U_{ij}, which we define as follows: If a respondent stated that she participated in on-the-job training once or

[4]The following version of the data set is used: Hall and Tiemann (2006) BIBB/BAuA Employment Survey of the Working Population on Qualification and Working Conditions in Germany 2006, SUF 1.0; Research Data Center at BIBB (ed.); GESIS Cologne, Germany (data access); Federal Institute of Vocational Education and Training, Bonn doi:10.4232/1.4820. For further details, also see Rohrbach (2009).

several times within the last two years or, alternatively, since being on her current job, we count either one as training incidence and set $U_{ij} = 1$. Otherwise we define $U_{ij} = 0$. The "BIBB/BAuA Employment Survey 2005/06" is particularly suited for our analysis since it combines detailed information on training participation with a rich set of individual controls that already have been identified as important determinants for the individual training decision (Bassanini, Booth, Brunello, De Paola, and Leuven, 2007). In particular, we have information on demographic controls (age, gender, education) and workplace characteristics (firm size, tenure, employment contract).[5] In context of the recent offshoring literature (cf. Acemoglu, Gancia, and Zilibotti, 2012), our data has the great advantage that we are able to observe the introduction of new technologies and organizational changes at the workplace. This allows us to discriminate between offshoring and technological change when explaining the variation in individual training decisions, and eliminates possible concerns about technological change being a potential source of an omitted variable bias. As another advantage of our data we have information on individual job loss fears (cf. Geishecker, Riedl, and Frijters, 2012). Given that offshoring often is associated with job losses for some workers (usually followed by a period of transitory unemployment and/or training) this information provides a suitable control for a potential postponement of on-the-job training in favour of later out-of-the-job training activity, as for example identified by Hummels, Munch, Skipper, and Xiang (2012). To control for business cycle effects, which have been linked to training by Méndez and Sepúlveda (2012), we rely on workers' assessment of the employing firm's current business success, but also compute industry level output growth between 2004 and 2006. Finally, following Görlitz and Stiebale (2011) we also use Herfindahl indices of industry concentration from the German Monopoly Commission for 2003 to control for varying product market competition in different industries.

[5] For sources, a comprehensive description, and more detailed summary statistics of the variables in our final sample please refer to the data appendix.

Offshoring is measured as a trade related phenomenon using data on imported intermediates.[6] In line with our identification strategy outlined above, we follow the literature and observe offshoring at the industry level (Ebenstein, Harrison, McMillan, and Phillips, 2013; Baumgarten, Geishecker, and Görg, 2013). In particular, we stick to the concept of Geishecker and Görg (2008) and use input-output tables provided by the German Statistical Office to compute the share Θ_{jj^*} of intermediate products used in industry j that originate from the same industry j^* abroad. We then multiply Θ_{jj^*} by IMP_j, which is the total value of sector j's imports of goods that originate from non-OECD countries, and finally divide by Y_j, which is the value of sector j's output. In the end we obtain

$$O_j = \frac{\Theta_{jj^*} IMP_j}{Y_j}, \tag{4.7}$$

as a measure for the intensity of offshoring in sector j. Note that our offshoring measure only includes intermediates that are imported from the same sector abroad, resembling the "narrow" concept of offshoring put forth in Feenstra and Hanson (1999).[7] Following our theoretical model from Section 4.1, we are interested in offshoring that results from a cost savings motive and, hence, focus only on imports of intermediates that originate from non-OECD countries.[8] After all, this gives us a measure of offshoring to non-OECD countries that varies across 22 manufacturing industries (according to the NACE 1.1 classification). We use this information to compute the sectoral growth rate of offshoring $\widehat{O}_j$ over the relevant sample period from 2004 to 2006. Both, levels and relative changes of our offshoring measure are reported in Table 8.2 (see Appendix B.1). The levels can be considered as fairly low, which reflects the fact that trade with non-OECD countries only accounts for a small share in German imports. Yet, growth has been

[6] Proxies for offshoring based on foreign direct investment (FDI) often suffer from the insufficient decomposability of this data with regard to the motive behind outbound foreign direct investments. As an exception in this literature, Davies and Desbordes (2012) are able to distinguish between greenfield FDI as well as mergers and acquisitions (M&A), which allows them to control for FDI motives such as technology acquisition or the elimination of foreign competitors.

[7] For a detailed discussion of the differences between the measure used here and the measure used by Feenstra and Hanson (1999) please refer to Geishecker and Görg (2008).

[8] See Grossman and Rossi-Hansberg (2012) a model of trade in tasks between similar countries, in which firms have incentives to cluster the production of the same tasks at the same location in the presence of external scale economies that operate at the country level.

impressive. On average offshoring increased by 33% over the period from 2004 through 2006. To obtain our final estimation sample, we match the growth rate of our offshoring variable with the individual information taken from the "BIBB/BAuA Employment Survey 2005/06" and our further sectoral control variables. Focusing only on individuals holding a full time contract in one of the 22 manufacturing industries considered above leaves us with a total of 3.917 observations.

4.2.3 Estimation results

We estimate several variants of the Probit model specified in Section 4.2.1. Starting with Table 4.1, in which we provide first evidence on the link between offshoring growth and on-the-job training, we gradually add additional individual control variables, which the training literature has identified as major determinants of individual skill upgrading (see Bassanini, Booth, Brunello, De Paola, and Leuven, 2007).

As a point of reference, Column (1) in Table 4.1 shows the average marginal effect of offshoring growth from 2004 to 2006 on the probability of on-the-job training participation. According to this first estimate, offshoring growth has a strong and significant impact on individual skill upgrading: A doubling of the non-OECD offshoring intensity defined in Eq. (4.7) would lead to an increase in the probability of on-the-job training participation by 0.1732. Taking into account the immense offshoring growth of (on average) more than 30% in the German manufacturing between 2004 and 2006, we find that a sizeable shift in training participation can be attributed to increased offshoring.

Gradually adding further individual controls in the Columns (2) to (6) downsizes the effect of offshoring growth only marginally. However, in line with Bassanini, Booth, Brunello, De Paola, and Leuven (2007) and Méndez and Sepúlveda (2012), we find the usual life-cycle pattern in the results in Column (2), according to which older individuals are less likely to undertake on-the-job training than their younger counterparts. Including a gender indicator in Column (3), we find, that men are more likely to select into on-the-job training than women, which at a first sight contrasts with the findings of Arulampalam, Booth, and Bryan (2004), who show that in the European context women are in general no less likely to participate in training than men. However, as documented in Bassanini, Booth, Brunello, De Paola, and Leuven (2007), the effect

Table 4.1: *Offshoring and on-the-job training: individual controls*

	(1)	(2)	(3)	(4)	(5)	(6)
Average marginal effect of:						
Offshoring growth	0.1732***	0.1643***	0.1570***	0.1565***	0.1549***	0.1500***
	(0.0534)	(0.0515)	(0.0490)	(0.0423)	(0.0415)	(0.0246)
Age 30 - 39		0.0351	0.0331	-0.0161	-0.0087	-0.0130
		(0.0228)	(0.0234)	(0.0254)	(0.0232)	(0.0199)
Age 40 - 49		-0.0142	-0.0132	-0.0855***	-0.0722**	-0.0691***
		(0.0280)	(0.0301)	(0.0290)	(0.0282)	(0.0242)
Age 50 - 64		-0.0964***	-0.0946***	-0.1970***	-0.1811***	-0.1725***
		(0.0330)	(0.0320)	(0.0280)	(0.0279)	(0.0247)
Age 65+		-0.3249***	-0.3257***	-0.4391***	-0.4214***	-0.4177***
		(0.0774)	(0.0788)	(0.0676)	(0.0665)	(0.0555)
female			-0.0630***	-0.0419**	-0.0393**	-0.0782***
			(0.0232)	(0.0200)	(0.0198)	(0.0176)
Married			-0.0100	-0.0148	-0.0147	-0.0112
			(0.0238)	(0.0235)	(0.0233)	(0.0226)
Tenure				0.0076**	0.0084**	0.0090**
				(0.0038)	(0.0039)	(0.0039)
Tenure squared				-0.0001	-0.0001	-0.0001
				(0.0001)	(0.0001)	(0.0001)
Medium-skill				0.1181***	0.1169***	0.0383
				(0.0372)	(0.0368)	(0.0350)
High-skill				0.2118***	0.2117***	0.0169
				(0.0255)	(0.0258)	(0.0204)
Importance to have a career					0.0628***	0.0651***
					(0.0199)	(0.0199)
KldB88 (2-digit) occupation FE	no	no	no	no	no	yes
Pseudo R-squared	0.0100	0.0199	0.0221	0.0492	0.0509	0.1133
Observations	3,917	3,917	3,917	3,917	3,917	3,888

Notes: The table shows average marginal effects from estimating variants of the Probit model specified in Section 4.2.1. The reference category for an individual's age is: age 16 - 29. Standard errors are clustered at the industry level and are shown in parentheses below the coefficients. Superscripts ***, **, and * denote statistical significance at the 1%, 5%, and 10% level, respectively.

of gender on training participation, crucially depends on the sector of employment, with woman receiving comparatively less on-the-job training in certain medium/low-tech manufacturing industries. Given that our sample only includes workers employed in manufacturing industries, with a strong bias towards male employment (on average 75.9%), we should not be surprised to find a negative gender coefficient. Marital status, which we also introduce in Column (3), has no significant effect on training participation. In Column (4) we additionally control for work experience and education. Tenure has a positive but small effect on the probability of training participation. We treat this result with caution, since tenure – for obvious reasons – most likely is endogenous (Bassanini, Booth, Brunello, De Paola, and Leuven, 2007). Turning to the education indicators, we find the usual result, that high-skilled workers are more likely to participate in training than medium-skilled workers, while medium-skilled workers are again more likely to participate in training than low-skilled workers (see Pischke, 2001; Bassanini, Booth, Brunello, De Paola, and Leuven, 2007). To control for usually unobservable heterogeneity among workers (e.g. motivation), we exploit the detailed information included in the "BIBB/BAuA Employment Survey 2005/06" and add a binary indicator variable, which takes the value of one if the individual stated that having a career is (very) important and a value of zero otherwise. As we would expect, individuals, which care more about their career, are also more likely to invest in individual skill upgrading. Finally, adding occupation fixed effects in Column (6) to account for occupation-specific variation in the data, leaves most of our coefficients unchanged.[9] Only the coefficients for education turn insignificant. This, however, does not come as a surprise, given that in Germany entry into most occupations is subject to strict skill requirements (e.g. holding a certain university degree or a specific vocational qualification). Taking into account the implied homogeneity of workers in terms of formal education within occupations, it is clear that any attempt to identify the education coefficients based on the remaining skill variation within occupations necessarily is doomed to fail. The necessity to control for occupation-specific effects in our context arises as interactivity and complexity in the job content of certain occupations impose severe limits to the offshorability of the respective jobs (Blinder, 2006; Goos, Manning, and Salomons, 2009; Ottaviano, Peri, and Wright, 2013). At the same time, these activities may require more frequent skill updating, which we would not want to confuse with our skill upgrad-

[9]By adding occupation fixed effects we lose 29 observations for which either no occupational classification is coded in the data or too few observation for the estimation of an occupation-specific effect exist.

ing mechanism from Section 4.1. Taking stock, we find that the effect of offshoring growth on on-the-job training participation is only marginally reduced if further control variables at the individual level are included.

In a next step we turn to more likely candidates for an omitted variable bias and control for characteristics, which either directly describe the individual workplace or link to the industry in which the respective worker is employed. We thereby keep our individual controls from Column (6) in Table 4.1 throughout, while gradually adding additional workplace- and industry-level control variables in Table 4.2.

We start with the inclusion of firm size controls in Column (1) of Table 4.2. In line with Bassanini, Booth, Brunello, De Paola, and Leuven (2007), we find that workers employed by larger firms are more likely to undertake on-the-job training than workers in small firms. Given that offshoring usually is highly concentrated among large firms, with small firms often doing no offshoring at all (see Moser, Urban, and Weder di Mauro, 2009; Hummels, Jørgensen, Munch, and Xiang, 2013), we would expect that our estimate is upward biased, if differences in firm size are not taken into account. Indeed, when controlling for differences in firm size, we find that the impact that offshoring growth has on the probability of individual skill upgrading is reduced, although still positive and highly significant. In Column (2) of Table 4.2 we add further controls, which directly describe the employees' individual working environments. In particular we take into account whether a worker is employed under a fixed term contract or through a temporary work agency. As in Arulampalam, Booth, and Bryan (2004) and Bassanini, Booth, Brunello, De Paola, and Leuven (2007), and in line with human capital theory, we find that workers employed under fixed term contracts are less likely to invest in skill acquisition than workers with permanent contracts. For workers temporary employed through an external supplier – after all only 1% of all workers in our sample – no such effect exists, which we attribute to a lack of variation in our data. Finally, we also take up recent findings by Geishecker, Riedl, and Frijters (2012), who claim that offshoring to low-wage countries can explain about 28% of the increase in subjective job loss fears of German workers for the time span from 1995 to 2006. Adding an indicator variable, which takes a value of one whenever individuals stated that they face the fear of job loss and zero otherwise, we find that workers who reported subjective job loss fears

Table 4.2: *Offshoring and on-the-job training: workplace and sectoral controls*

	(1)	(2)	(3)	(4)	(5)	(6)
Average marginal effect of:						
Offshoring growth	0.1110***	0.1107***	0.1060***	0.1078***	0.1043***	0.0776***
	(0.0253)	(0.0253)	(0.0253)	(0.0257)	(0.0264)	(0.0201)
Firm size 10 - 49	-0.0036	-0.0006	-0.0148	-0.0191	-0.0147	-0.0187
	(0.0225)	(0.0216)	(0.0216)	(0.0218)	(0.0216)	(0.0214)
Firm size 50 - 249	0.0701***	0.0759***	0.0537***	0.0488***	0.0537***	0.0496***
	(0.0170)	(0.0150)	(0.0162)	(0.0166)	(0.0161)	(0.0163)
Firm size 250 - 499	0.1241***	0.1306***	0.1051***	0.0990***	0.1049***	0.0973***
	(0.0303)	(0.0288)	(0.0278)	(0.0272)	(0.0279)	(0.0281)
Firm size 500+	0.1518***	0.1584***	0.1343***	0.1273***	0.1343***	0.1190***
	(0.0284)	(0.0273)	(0.0267)	(0.0253)	(0.0266)	(0.0243)
Fixed term contract		-0.0901***	-0.0730**	-0.0765**	-0.0730**	-0.0785**
		(0.0323)	(0.0318)	(0.0320)	(0.0319)	(0.0328)
Temporary work		0.0272	0.0557	0.0512	0.0571	0.0394
		(0.0532)	(0.0538)	(0.0532)	(0.0535)	(0.0539)
Job loss fear		-0.0621***	-0.0632***	-0.0470**	-0.0634***	-0.0504**
		(0.0204)	(0.0208)	(0.0207)	(0.0209)	(0.0211)
New technology introduced			0.1674***	0.1655***	0.1676***	0.1640***
			(0.0219)	(0.0218)	(0.0218)	(0.0214)
Current Firm success (very) good				0.0429**		0.0406**
				(0.0200)		(0.0192)
Industry level output growth					-0.0614	
					(0.0993)	
Industry level Herfindahl index						0.0006***
						(0.0001)
Individual controls	yes	yes	yes	yes	yes	yes
KldB88 (2-digit) occupation FE	yes	yes	yes	yes	yes	yes
Pseudo R-squared	0.1240	0.1271	0.1359	0.1369	0.1360	0.1391
Observations	3888	3888	3888	3888	3888	3888

Notes: The table shows average marginal effects from estimating variants of the Probit model specified in Section 4.2.1. The reference category for firm size is 1 - 9 employees. The industry output growth is computed for 2004 to 2006. The Herfindahl index, which is published bi-annually by the German Monopoly Commission refers to 2005. Individual controls are the same as in Column (6) of Table 4.1. Standard errors are clustered at the industry level and shown in parentheses below the coefficients. Superscripts ***, **, and * denote statistical significance at the 1%, 5%, and 10% level, respectively.

are less likely to invest in on-the-job training. Together with the findings of Hummels, Munch, Skipper, and Xiang (2012), who show that workers who lose their job (through offshoring) are more likely to retrain their skills during the subsequent period of transitory unemployment, this result may hint at a delay of on-the-job training in favour of later out-of-job training measures, which are better tailored towards future re-employment possibilities. Important in our context is that none of these controls do significantly alter the average marginal effect of offshoring growth on individual skill upgrading. We now turn to Column (3) of Table 4.2, in which we include a binary variable that takes a value of one whenever new technologies, machines, or organizational features have been introduced at individual workplaces. There are two specific reasons why we have to control for the introduction of new technologies in our setting: On the one hand, our theoretical model from Section 4.1 reveals a close resemblance between the productivity effect of offshoring and sector biased technological change, which we have to tell apart if we want to identify the impact of offshoring growth on individual skill upgrading (cf. Feenstra and Hanson, 1999; Feenstra, 2010). On the other hand, it is likely that whenever new technologies are introduced this requires the (re-)training of involved workers, thereby mechanically leading to increased on-the-job training, which we do not want to confuse with our skill upgrading channel from Section 4.1. In line with these arguments, we find that workers who reported the introduction of new technologies at their workplace are more likely to participate in on-the-job training. Crucially, there still is a positive and highly significant link between offshoring growth and individual skill upgrading, although – as we would expect – with a lower estimate of the average marginal effect, which now stands at $\hat{\beta}^m = 0.1060$. A further concern relates to a possible co-movement of increased offshoring with the sectoral business cycle. If on-the-job training is pro-cyclical, for which – despite partly confounding results – at least some evidence exists (cf. Méndez and Sepúlveda, 2012), it could be the case that the positive association of individual skill upgrading with increased offshoring is nothing else than the reflection of the German business cycle, which from 2004 to 2006 was at the beginning of a boom period. To rule out this possibility, we include in Column (4) of Table 4.2 a control variable, which reflects workers' evaluation of the employing firms' current business success. In line with Méndez and Sepúlveda (2012), we find that workers employed by (very) successful firms tend to invest more often in on-the-job training. At the same time, the effect of offshoring growth on skill upgrading is almost unchanged. Admittedly, our measure for the business cycle is a simple one, focusing only on the employing firm, thereby ignoring possible inter-firm linkages in the respective industry. To come up with more comprehensive measure we also add the log-difference of real industry output in Column (5) of Table 4.2.[10] The effect of output growth on on-the-job training is insignificant, which is in line with the somewhat inconclusive literature on the cyclical properties of training (see Méndez and Sepúlveda, 2012; Bassanini, Booth, Brunello, De Paola, and Leuven, 2007). Not surprisingly, the effect of offshoring on skill upgrading is only slightly reduced and stays highly significant. Finally, in Column (6) of Table 4.2 we also control for the competition intensity within a given sector (cf. Görlitz and Stiebale, 2011). Given the positive correlation between firm size and offshoring activities, it could be the case that industries dominated by a few large firms have significantly different offshoring growth patterns than industries which are characterised by a competitive number of firms. At the same time skill upgrading – for several reasons – may also be linked to the intensity of competition within a sector: On the one hand increased competition could lead to higher training needs, necessary to secure a well trained workforce in a dynamic environment (Bassanini and Brunello, 2011). On the other hand, poaching, i.e. the transfer of general skills to a different employer via job switching, is usually found to be positively correlated with competition, which, hence, would lead to less training (Schmutzler and Gersbach, 2012). Controlling for industry level competition, we use the same measure as Görlitz and Stiebale (2011), the Herfindahl index of industry concentration.[11] We find a positive impact of competition on training, which is significant at the 1% level. Importantly, the effect of offshoring growth on individual skill upgrading is still significant, albeit slightly smaller in magnitude. Summing up, we find, that according to our preferred specification in Column (6) of Table 4.2 a doubling of the industry level offshoring intensity defined in Eq. (4.7) would increase the probability of on-the-job training participation by roughly 7.8 percentage points.

[10]Including sectoral output growth might raise concern about possibly high multicollinearity between output growth and offshoring growth. However, this does not seem to be the case, as the coefficient on output growth stays insignificant even if offshoring growth is excluded from the regression.

[11]The Herfindahl index is published bi-annually by the German Monopoly Commission. We use the values for 2005.

4.2.4 The timing of offshoring and on-the-job training

The "BIBB/BAuA Employment Survey 2005/06" took place from October 2005 and March 2006, and individuals where asked whether they participated in on-the-job training two years prior to the survey or since having the current job. We, thus, hold no precise information concerning the timing of individual training incidences. For our main analysis we therefore use a rather wide time frame covering offshoring growth from 2004 to 2006. As a robustness check, we now consider shorter and varying time frames, each covering the growth rate of offshoring on a year-to-year basis. Results are summarized in Table 4.3.

As we would expect, splitting up the time frame from 2004 to 2006 into two separate windows, covering 2004 to 2005 and 2005 to 2006, does not change our result: Increased offshoring still has a positive and significant impact on individual skill upgrading. Looking at a period (2002 - 2003) that far precedes the time frame potentially covered by the survey we find – as expected – no effect. For the time frame from 2003 to 2004 we find a negative and significant coefficient. We interpret this finding as evidence, that on-the-job training is a lumpy investment, which individuals only use discontinuously over time with an optimal period of waiting between single training incidences. Thus, if increased offshoring between 2003 and 2004 caused more training in the period from 2003 to 2004 we would indeed expect that in the following period from 2004 to 2006 an immediate retraining becomes less likely for individuals who just completed their last training in the previous period. Interestingly, when focusing on future offshoring growth over the period from 2006 to 2007 we find a negative impact on the current training probability. Several explanations may account for this result. Assuming that individuals discount the costs and benefits of training at different rates, it could be the case that anticipated future offshoring growth leads to a postponement of contemporaneous training to later period when the benefits from training are even larger. Another explanation for the negative impact of future offshoring growth on contemporaneous training could be the lack of sufficient *long-run* controls capturing individual job loss fears. Given that offshoring tends to increase subjective job loss fears (cf. Geishecker, Riedl, and Frijters, 2012), anticipated future offshoring growth could be associated with more uncertain long-run employment prospects, causing a reduction or delay of current on-the-job training. Finally, in Column (6) of Table 4.3 we only use individuals that are em-

Table 4.3: *Offshoring and on-the-job training: timing*

	(1)	(2)	(3)	(4)	(5)	(6)
Average marginal effect of:						
Offshoring growth 2002 - 2003	0.0178					
	(0.0121)					
Offshoring growth 2003 - 2004		-0.1589**				
		(0.0697)				
Offshoring growth 2004 - 2005			0.2672***			
			(0.0536)			
Offshoring growth 2005 - 2006				0.2216***		
				(0.0774)		
Offshoring growth 2006 - 2007					-0.1340***	
					(0.0476)	
Offshoring growth 2004 - 2006						0.1009***
						(0.0252)
Individual controls	yes	yes	yes	yes	yes	yes
Workplace and sectoral controls	yes	yes	yes	yes	yes	yes
KldB88 (2-digit) occupation FE	yes	yes	yes	yes	yes	yes
Pseudo R-squared	0.1335	0.1350	0.1370	0.1362	0.1354	0.1490
Observations	3,888	3,888	3,888	3,888	3,888	3,425

Notes: The table shows average marginal effect from estimating the variants of the Probit model in section 4.2.1 for different periods of offshoring growth. The dependent variable is a binary measure of observed skill upgrading through training in the two years prior to the survey or since having the current job. The Herfindahl index is not included since we do not have it for all respective time periods. Standard errors are cluster robust and are shown in parentheses below the coefficients. Superscripts ***, **, and * denote statistical significance at the 1%, 5%, and 10% level, respectively.

ployed at the same employer since at least 2003. Because individuals where asked whether they participated in on-the-job training two years prior to the survey or since having the current job, this treatment gives us a more precise matching of the potential training period with the time frame for which we observe our offshoring variable. The resulting coefficient for sectoral offshoring growth is very similar to the one obtained from our preferred specification (Column (6) of Table 4.2) and highly significant.

4.2.5 Further robustness checks

In this subsection, we offer alternative specifications and check whether the link between offshoring growth and skill upgrading is driven by particular characteristics of our data set in terms of measurement or outliers. Detailed results can be found in Table 8.3 in the Appendix. At first, in Column (1) of Table 8.3 we look at the growth rate of worldwide offshoring, instead of the growth rate of non-OECD offshoring intensities. We do this to provide evidence for an alternative measure of offshoring. We find a positive and significant coefficient – which, somewhat surprisingly, is even larger than what we have estimated before. Secondly, in Column (2) we look at non-OECD offshoring again and include the growth rate of the export share in production to control for influences related to overall international exposure. In column (3) we use sample weights provided in the data and re-run our preferred specification using these weights. Note, however, that the data set is designed by the BIBB to be balanced and adjustments are taken to control for under representation of low-skilled individuals. Thus, it is not surprising to observe very similar coefficients, both in terms of significance and magnitude. Next, we drop in Column (4) four industries (tobacco; leather & luggage; office machinery & computers; coke & refined petroleum) in which results, due to a low number of observations, could easily be affected by outliers. In Column (5) we drop the two industries with the largest (other transport equipment) and smallest (coke & refined petroleum) change in non-OECD offshoring, again to rule out dependence on outliers, which could play an important role in our relatively small sample. Similarly, in Column (6) we drop the industries with the highest (chemicals) and lowest (textiles) average training participation rates. Reassuringly, all those changes have almost no effect on the coefficient of sectoral offshoring growth, which remains positive and significant throughout all specifications. Finally, let us recall our theoretical model from section 4.1, in which training participation is modeled as a worker's decision and it is the worker to whom both the cost and the benefits associated with individual skill upgrading accrue. Translating this mechanism one to one into our empirical model would require a distinction between employer-financed and self-financed on-the-job training. Unfortunately, this information is not available to us. However, we know whether a certain training measure can be traced back to the respective worker's own initiative or to some extrinsic motivation. Assuming that training which workers' started by own initiative is more likely to be also self-financed, we drop all cases in which workers' training participation followed from the order or suggestion of the respective employer. The results are shown in Column (7). Controlling for workers' initiative to start on-the-job training does not imply a correction of effect the growth rate of offshoring has on individual training participation. Importantly, the coefficient is still significant and of similar size when compared to the coefficient that results from the estimation of the full sample.

4.3 Summary

In this study we have derived a positive link between the offshoring of tasks and the individual propensity to invest in on-the-job training. In particular, we developed a theory that outlines a mechanism inducing employed individuals to select into training – a new aspect in the literature linking offshoring and training, which has so far mostly analysed training responses to worker displacement. In our model offshoring allows firms to save on costs when relocating parts of their production abroad. The resulting costs savings are handed through to domestic workers, whose wages are scaled up, thereby opening up so far unrealized skill upgrading possibilities. We test for this intuitive mechanism, using data from German manufacturing, and find that industry level offshoring growth rates indeed correlate with the individual probability of on-the-job training in a positive and highly significant way. In obtaining this effect, we explicitly control for a wide set of individual and workplace characteristics and in particular take into account technological change at the workplace and industry output growth as major determinants of individual on-the-job training.

Chapter 5

Two-way migration between similar countries

5.1 Permanent migration

In this section we develop a simple model to explain permanent two-way migration of high-skilled individuals between developed countries. While the phenomenon of two-way migration has received little attention in the theoretical literature, it is quantitatively important, in particular for high-skilled individuals migrating between high-income countries. Table 5.1, which is based on data from Docquier, Lowell, and Marfouk (2008), shows for country pairs within the EU15 and the OECD, respectively, the share of bilateral migration that can be characterised as two-way. The share is measured by the Index of Bilateral Balance in Migration (Biswas and McHardy, 2005), which for each country pair (i, j) is given by $B_{ij} \equiv 2 \min(\mathrm{Em}_{ij}, \mathrm{Em}_{ji})/(\mathrm{Em}_{ij} + \mathrm{Em}_{ji})$, with Em_{ij} as the stock of emigrants from country i residing in country j.[1] The numbers in Table 5.1 are the average values of the index for the respective country group, in a given year and skill group. The data show that the share of two-way migration is highest for high-skill individuals, that it has grown over time, and that it is higher within the homogeneous group of EU15 countries than in the more heterogeneous group of OECD countries.

[1]The construction of the index is directly analogous to the well-known Grubel-Lloyd index measuring intra-industry trade, i.e. two-way trade in goods within the same industry.

Table 5.1: *Index of bilateral balance in migration for EU15 and OECD countries*

		high skill	med. skill	low skill	total
EU15	1990	0.61	0.53	0.20	0.35
	2000	0.64	0.51	0.28	0.48
OECD	1990	0.23	0.22	0.14	0.19
	2000	0.28	0.26	0.13	0.22

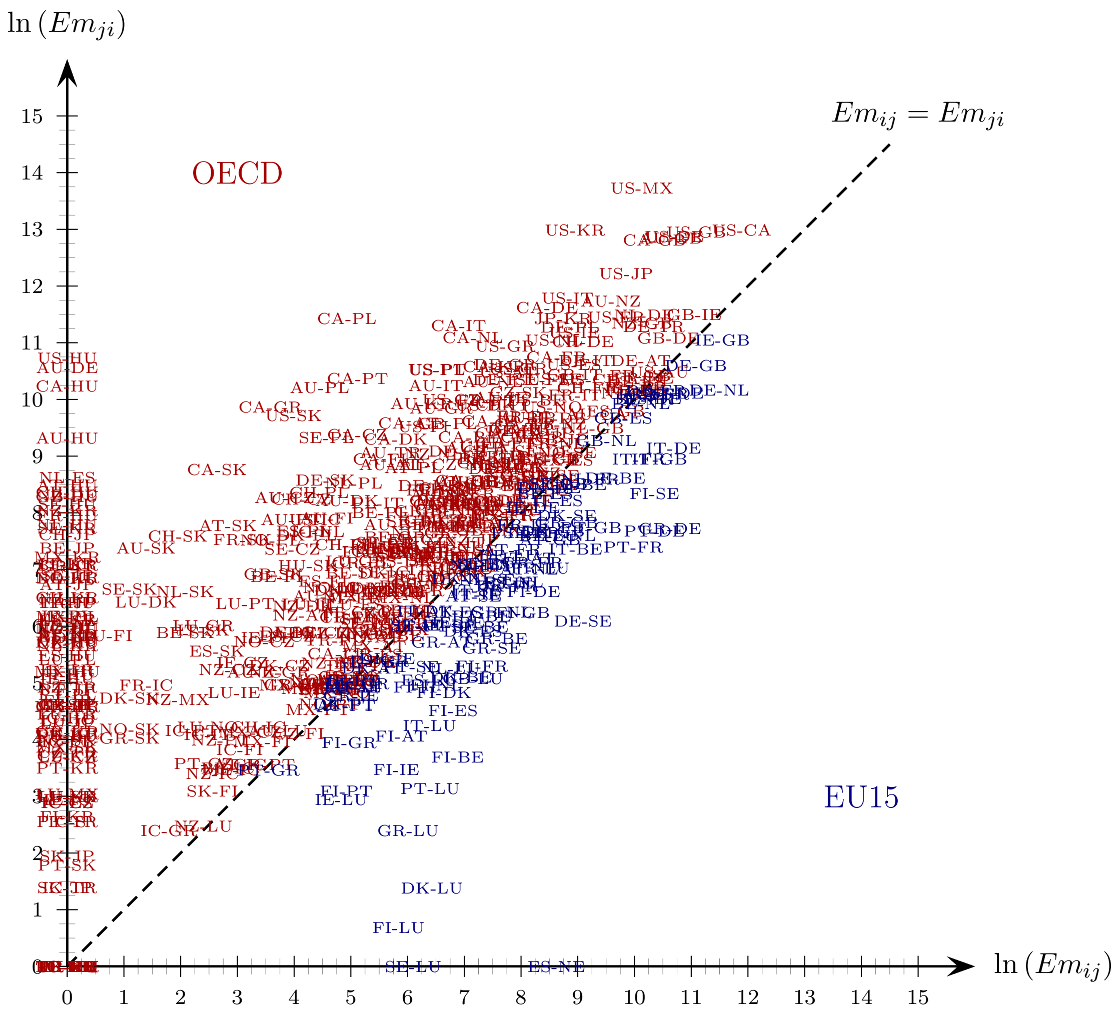
Focussing on high-skilled (tertiary educated) individuals, Figure 5.1 gives a more disaggregated view at the level of country pairs for the EU15 (below the diagonal) and the OECD (above the diagonal).[2] The figure confirms that a lot of high-skilled migration between EU15 countries is two-way in nature, while this is true to a lesser extent for the larger and more heterogeneous group of OECD countries. Despite this regularity there is, of course, incidence of substantial two-way migration for specific country pairs that are part of the OECD but not part of the EU15. Taking Canada and the US as another prominent example of rather similar countries, we observe substantial high-skilled migration in both directions, with the share of two-way migration being 0.5 for the year 2000.[3]

The key challenge in explaining two-way migration of similar (highly skilled) individuals within a group of similar (high-income) countries – rather than one-way migration from low-income to high-income countries – lies in the fact that country differences cannot be expected to play a central role. The model we develop in this section therefore uses the assumption that

[2]The figure plots $(15\times14)/2 = 105$ country pairs from the set of EU15 countries and $(30\times29)/2 = 435$ country pairs from the sample of OECD countries. Chile, Estonia, Israel and Slovenia are omitted, as data on two-way migration is not available for these countries in Docquier, Lowell, and Marfouk (2008). Note that country pairs are order such that for the set of EU15 countries (blue) the net-emigration country appears first, while for the set of OECD countries (red) the net-immigration country is named first. Hence, the strict separation in above and below the 45 degree line.

[3]See Schmitt and Soubeyran (2006) and the references cited therein for additional anecdotal evidence on the balance in migration flows between Canada and the US.

Figure 5.1: *Two-way migration among EU15 and among OECD countries*



countries are identical in all respects (this assumption is relaxed later on). In both countries there is a continuum of workers with differing abilities, which are private knowledge. The production technology, borrowed from Kremer (1993), exhibits complementarities between the skill levels of individual workers, and profit maximising firms therefore aim for hiring workers of identical skill. Migration is costly, and the cost is the same for all individuals. High-skilled individuals from both countries self-select into emigration in order to separate themselves from low-skilled co-workers at home. Firms can distinguish natives and immigrants, which allows them to form more efficient matches, leading to larger gross wage premia for skilled workers.

The welfare effects of migration in our model are stark: In the laissez-faire equilibrium all individuals are worse off than in autarky. We show that this result is due to a negative migration externality which leads to too much migration in equilibrium. We also show that for sufficiently low migration cost the level of migration chosen by an omniscient social planner is strictly positive (but of course lower than in the laissez-faire equilibrium), since the existence of migrants as a distinct group of individuals enables firms to match workers of more similar expected skill. While aggregate gains from migration exist in the social planner equilibrium, the distributional effects are strong: All migrants gain relative to autarky, while all natives are worse off. These distributional effects are mitigated if the social optimum is implemented via a migration tax, since in this case the possibility of redistributing part of the gains to non-migrants exists.

Our baseline model is deliberately stylised in order to bring out the basic mechanism driving two-way migration and its welfare implications in the most transparent way possible. Due to its simplicity, the basic version of the model has some extreme implications, and we introduce multiple extensions with the aim for the model to better replicate various stylised facts of international migration. In a first extension, we consider a situation where skills are imperfectly observable, rather than unobservable as in our benchmark model. The most important effect of this change is to give rise to instances where firms co-hire migrants and natives, thereby mitigating the complete segregation between migrants and natives across firms that is implied by our basic framework (cf. Hellerstein and Neumark, 2008; Andersson, García-Pérez, Haltiwanger, McCue, and Sanders, 2010). Our second extension analyses two-way migration in a world where skills are only imperfectly transferable across countries (cf. Mattoo, Neagu, and Özden, 2008; Chiswick and Miller, 2009). In this extension the skill distributions of migrants and natives overlap, giving rise to a scenario in which migrants can find themselves in the middle (instead of on the top) of the destination country's skill distribution. Acknowledging that migration in our framework effectively acts as a signalling device, we then show that migration is still observed as an equilibrium phenomenon in our model if we add alternative signals, as for example education (cf. Spence, 1973). In a fourth extension of our model we add capital as an internationally immobile factor that is an essential input in the production of all firms (cf. Kremer, 1993). This extension introduces into our framework interactions between migrants and domestic factors of production, which are well known in principle from many existing models of international migration (see, e.g., the complementarity between labour and capital underlying the "immigration surplus" first documented in Berry and Soligo (1969) and more recently reviewed by Borjas (1999), or the imperfect substitutability between natives and migrants recently highlighted in Ottaviano and Peri (2012)). We show that migration is potentially more benign in this case than in our basic model, since it allows for the more efficient allocation of capital between domestic firms, with firms hiring migrants having a higher capital intensity due to a capital-skill complementary that is well known from many models of migration. Lastly, we allow for small differences in countries' technologies. By gaining access to a better technology, workers from the low-tech country then have an additional incentive to migrate, while the opposite holds true for workers from the high-tech economy. Incorporating this modified incentive structure, we still find two-way migration, which now is, however, biased towards the technologically superior country: The high-tech country experiences net immigration while the low-tech country faces net emigration.

The vast majority of theoretical models on high-skilled migration are in the tradition of the "brain drain" literature, focussing on high-skilled migration from developing to more advanced economies. Early contributions to this literature focused on the economic losses for source countries.[4] However, more recently the possibility of a net "brain gain" as the prospect of emigration raises education incentives has been emphasised by Mountford (1997), Stark, Helmenstein, and Prskawetz (1997) and Beine, Docquier, and Rapoport (2001).[5] Embedding high-skilled migration between asymmetric countries into a general equilibrium model of inter- and intra-industry trade Iranzo and Peri (2009) show that source countries gain, if high-skilled migration and trade are complements and gains from trade through a larger set of varieties accrue globally. Similarly, Bougheas and Nelson (2012) find that the majority of workers in source and sending countries benefit from high-skilled emigration as Ricardian-type comparative advantages and the gains

[4] Grubel and Scott (1966b) point to the loss of positive externalities as professionals emigrate. Bhagwati and Hamada (1974) stress the fiscal loss associated with the emigration of high-income earners, while Wong and Yip (1999) show that a brain drain has negative growth effects as human capital accumulation is deteriorated.

[5] For a detailed review of the brain drain/gain literature see Hanson (2010) or Docquier and Rapoport (2012).

from trade associated with it are reinforced.[6] Hendricks (2001) use the same basic migration mechanism as we do and models costly emigration as a signalling device, which is used by the most able individuals to reveal their high but otherwise unobservable skills.[7] Unlike our paper, which analyses two-way migration between similar countries, Hendricks (2001) thereby focuses on one-way migration and the subsequent assimilation of migrants into the more advanced destination economy.

What all these models have in common are directed flows of high-skilled migrants from less to more advanced economies triggered by exogenous country asymmetries. To the best of our knowledge, we are the first to develop a model that can explain two-way international migration of high-skilled workers between identical countries. Schmitt and Soubeyran (2006) address the interesting but distinct question of two-way migration by individuals that have the same occupation, rather than the same skill level. In their model, individuals have either high skills or low skills, and they choose to be either entrepreneurs or workers, as in Lucas (1978). The career choice of individuals depends not only on their own skill level, but also on the skill distribution within each country. The equilibrium may feature two-way migration of both entrepreneurs and workers, but high-skill individuals only migrate to the country where skills are relatively scarce. If the countries are identical, as assumed in the main part of our model, no migration occurs. Our model is also related to Hendricks (2001) and Giannetti (2001), who use the same basic selection mechanism of high-skilled individuals into emigration as we do. But neither of these models analyses two-way migration, which is the question we focus on in this chapter.

5.2 A simple model of permanent migration

Consider a world with two perfectly symmetric countries, each populated by a heterogeneous mass of workers, which we normalise to one without loss of generality.[8] Workers in each country

[6] For a discussion of the complementarity between international migration and international trade see for example Felbermayr, Grossmann, and Kohler (2012).

[7] See also Giannetti (2001), who also models migration as a signalling device to explain inter-regional migration patterns in Italy.

[8] Since countries are assumed to be symmetric, we suppress all country indices.

differ with respect to their skills s which are uniformly distributed over the interval $[0,1]$, and which are assumed to be private information. Moreover, workers are risk neutral, such that utility $u(x) = x$ can be expressed as a linear function of consumption x. Each country is a single sector economy producing a homogeneous *numéraire* good y under perfect competition, which is costlessly traded.

We follow Kremer (1993) in assuming a production technology which requires the processing of $l = 1, 2$ tasks, each to be performed by a single worker. Firm-level output is given by

$$y = f(s_1, s_2) = 2As_1s_2, \tag{5.1}$$

where $A > 0$ is a technology parameter and s_l denotes the skill level of a worker performing task $l = 1, 2$. Note that $\partial f(s_1, s_2)/\partial s_l > 0 \;\forall\; l = 1, 2$. Moreover we have $\partial^2 f(s_1, s_2)/\partial s_l \partial s_{\hat{l}} > 0$ for all $l, \hat{l} = 1, 2$ and $l \neq \hat{l}$, such that Eq. (5.1) is supermodular and workers enter production as complements.

In an equilibrium that features migration, firms can identify an individual worker as a member of either the group of natives or the group of immigrants. This is the only information they can base their hiring decision on, and this information is valuable since, as we show below, the average skill of the two groups is different. Firms maximise their expected profits by choosing the optimal skill mix of their employees:

$$\max_{\bar{s}_1, \bar{s}_2} \pi(\bar{s}_1, \bar{s}_2) = 2A\bar{s}_1\bar{s}_2 - w(\bar{s}_1) - w(\bar{s}_2), \tag{5.2}$$

with $\bar{s}_l$, $l = 1, 2$, denoting the average skill of the group from which the worker for task l is hired, and $w(\bar{s}_l)$ being the expected wage paid to this worker. Lemma 5.2.1 gives the solution to this optimisation problem.

Lemma 5.2.1 *Firms maximise expected profits by hiring workers of the same expected skill.*

Proof See Appendix A.18.

Wages cannot be based on individual ability, since this is private information. Consequently, each worker is paid half the firm's output independent of her actual contribution. Using this remuneration rule, the expected wage rate of an individual worker with skill s equals

$$w(\bar{s}_\ell, s) = A\bar{s}_\ell s, \tag{5.3}$$

where $\bar{s}_\ell$ with $\ell \in \{L, H\}$ is the average skill of the group to which the individual belongs. We assume that migration is costly, and the cost is equal to c. Although workers cannot observe the individual skill of their potential co-workers, the distribution of skills in both countries is known, such that expectations can be formed with regard to a potential co-worker's average skill $\bar{s}_\ell$. It is now straightforward to show that our model leads to self-selection of the most able individuals into emigration.

To see this, consider some arbitrary cutoff ability $\tilde{s}$ that separates high-skill and low-skill individuals. The average skills in the two groups, L and H, are $\bar{s}_L = \tilde{s}/2$ and $\bar{s}_H = (1+\tilde{s})/2$ due to our assumption of a uniform distribution, and the resulting difference between the averages of both groups $\bar{s}_H - \bar{s}_L$ is equal to $1/2$ for all values of $\tilde{s}$. The expected wage gain for an individual worker of being paired with a co-worker from group H is now given by $A(\bar{s}_H - \bar{s}_L)s = As/2$, and it follows immediately that this gain is increasing in an individual's skill s. With identical migration cost for each individual, and assuming an interior solution, i.e. $0 < \tilde{s} < 1$, it follows that high-skilled individuals self-select into migrating abroad, while low-skilled individuals are deterred from migration by the cost attached to it. For the indifferent worker with skill $\tilde{s}$ the condition $A\tilde{s}/2 = c$ holds, which immediately gives the migration cutoff in the *laissez-faire* equilibrium as

$$\tilde{s}^{\mathrm{lf}} = \frac{2c}{A}. \tag{5.4}$$

Self-selection into migration with $\tilde{s}^{\mathrm{lf}} \in (0,1)$ then obviously requires $c \in (0, A/2)$. Proposition 5.2.2 summarises:

Proposition 5.2.2 *With strictly positive but not prohibitively high migration cost, all workers with skill $s > \tilde{s}^{lf} = 2c/A$ emigrate, while all workers with skill $s \leq \tilde{s}^{lf} = 2c/A$ stay in their home country. Migration flows increase for a higher level of technology A, and for lower migration cost c.*

Proof See Appendix A.19.

Taking stock, our model is able to explain two-way, high-skilled migration flows between two ex ante and ex post symmetric countries, which are driven by the desire of high-skilled workers to get separated from their low-skilled counterparts. In the resulting equilibrium costly migration acts as a signalling device, allowing high-skilled workers to (partly) reveal their true skill levels as in Spence (1973).

Lemma 5.2.1 and Proposition 5.2.2 together imply that firms hire only migrants or only natives. While this extreme implication of our model is counterfactual of course, Hellerstein and Neumark (2008), Andersson, García-Pérez, Haltiwanger, McCue, and Sanders (2010), Aslund and Skans (2010), Dustmann, Glitz, and Schönberg (2011) and Glitz (2012) find that there is indeed considerable segregation of natives and migrants across workplaces in the US, Sweden and Germany.[9] Interestingly, Hellerstein and Neumark (2008), Andersson, García-Pérez, Haltiwanger, McCue, and Sanders (2010) and Aslund and Skans (2010) also find that the high degree of workplace segregation between natives and migrants in the US and Sweden is only weakly related to the workers' general education. Picking up on this, we show in Section 5.4.1 below that a simple extension of our model, in which the abilities of some individuals are observable, is compatible with the empirical observation of imperfect workplace segregation between (high-skilled) natives and migrants.

As another straightforward implication of Proposition 5.2.2 we find the extreme result that within each country migrants are at the top, while natives are at the bottom of the skill distribution with no overlap in both groups' skill ranges. Modelling imperfectly transferable skills in line with the empirical findings by Mattoo, Neagu, and Özden (2008) and Chiswick and Miller (2009) we show in Section 5.4.2 that our model can account for an overlap in the skill range of migrants and natives. Alternatively we show in Section 5.4.3 that a similar result can be obtained if workers can choose between migration and education as signalling devices. Since the key mechanisms driving migration in our model are unaffected by these extensions, we stick to our more parsimonious formulation with unobservable but perfectly transferable skills and migration as the only signalling device for the time being, in order to save on notation and

[9] Hellerstein and Neumark (2008) find that 39.4% of Hispanics in the US have a co-worker who is also Hispanic, while only 4.5% of the white workers have Hispanic co-workers. Comparing this to a probability of 6.9% for having a Hispanic co-worker under random matching reveals a substantial workplace segregation by ethnicity. Figure 1 in Andersson, García-Pérez, Haltiwanger, McCue, and Sanders (2010) plots the cumulative distribution of the immigrant co-worker share for natives and migrants, respectively, which significantly differ from the distribution that would result under random assignment.

terminology.

5.3 Welfare effects of permanent migration

In order to analyse the welfare effects of migration, the natural comparison is the scenario of prohibitive migration cost $c \geq A/2$, which leads to $\tilde{s}^{\mathrm{lf}} = 1$ (the "autarky case"). The value of aggregate production equals total wage income, which for an arbitrary cutoff $\tilde{s}$ is given by

$$Y(\tilde{s}) = A\left[\int_0^{\tilde{s}} \left(\frac{\tilde{s}}{2}\right)^2 \mathrm{d}s + \int_{\tilde{s}}^1 \left(\frac{1+\tilde{s}}{2}\right)^2 \mathrm{d}s\right] = \frac{A\left[1+\tilde{s}(1-\tilde{s})\right]}{4}.$$

Total output is therefore minimised under autarky ($\tilde{s} = 1$), and maximised if exactly half the individuals become migrants ($\tilde{s} = 1/2$). Aggregate output rises, since firms that recruit their workers from a labour market with a more diverse labour supply are able to discriminate between the groups of natives and migrants. Since workers of the same nationality are more similar with respect to their (unobservable) skills, we find, that firms in a more fractionalised labour market realise productivity gains (cf. Trax, Brunow, and Suedekum, 2012) from the more efficient matching of workers according to Lemma 5.2.1.

Aggregate welfare equals the difference between total output and total migration cost:

$$W(\tilde{s}, c) = \frac{A[1+\tilde{s}(1-\tilde{s})]}{4} - c(1-\tilde{s}). \tag{5.5}$$

We can now use the link between $\tilde{s}^{\mathrm{lf}}$ and c provided by Eq. (5.4) to express aggregate welfare in the laissez-faire equilibrium as a function of $\tilde{s}^{\mathrm{lf}}$ alone:

$$W(\tilde{s}^{\mathrm{lf}}) = \frac{A\left[1-\tilde{s}^{\mathrm{lf}}(1-\tilde{s}^{\mathrm{lf}})\right]}{4}.$$

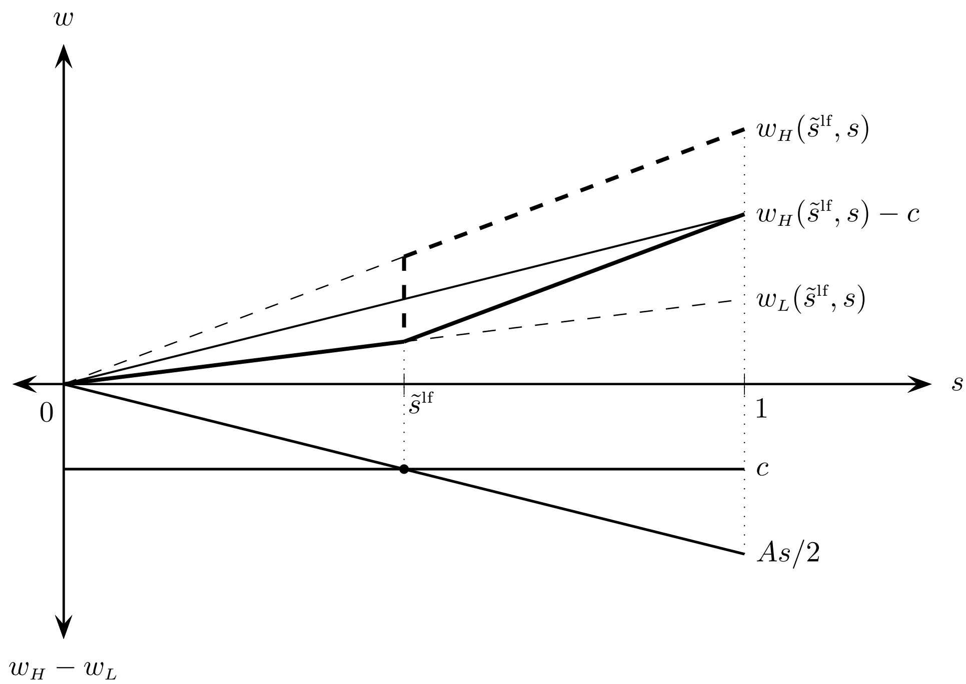
Thus, the effect of migration on aggregate welfare is diametrically opposed to its effect on total output: Aggregate welfare is maximised under autarky ($\tilde{s}^{\mathrm{lf}} = 1$), and minimised if exactly half the individuals become migrants ($\tilde{s}^{\mathrm{lf}} = 1/2$).

We now look at individual welfare, which is identical to an individual's expected wage rate, net of migration cost, if applicable. Non-migrants' and migrants' welfare is given by

$$w_L(\tilde{s}^{\mathrm{lf}}, s) = \frac{A\tilde{s}^{\mathrm{lf}}s}{2} \qquad \text{and} \qquad w_H(\tilde{s}^{\mathrm{lf}}, s) - c = \frac{A[s-\tilde{s}^{\mathrm{lf}}(1-s)]}{2},$$

respectively. We see that all individuals are worse off than in the autarky equilibrium, where the expected wage rate of an individual with skill s is equal to $As/2$.[10] For non-migrants, this simply happens because the pool of co-workers available for matching now has a lower average skill. For migrants, this is explained by a negative external effect induced by migration that can best be seen by a thought experiment, in which individual migration occurs sequentially, in the order of decreasing ability of migrants: Every migrant, apart from the most skilled one, in this case reduces the average skill of individuals in the migrant pool, thereby inflicting losses on infra-marginal migrants' wages. This effect is rationally ignored by individual migrants. Figure

Figure 5.2: *Laissez-faire equilibrium*



5.2 illustrates the results. The bottom quadrant shows how the migration cutoff is determined by the equality of migration cost and expected migration gain for the marginal migrant. The top quadrant shows in bold the resulting wage profile in the open economy as a function of individual

[10]Of course this result depends on the assumed skill distribution. As we show in Appendix A.20, an equilibrium in which every worker is worse off results for all skill distributions, which feature a convex cumulative density function, while for skill distributions with concave cumulative density functions there are net gains from migration for the most able migrants.

ability s where for migrants a distinction is made between the gross wage (bold dashed) and the net wage, which subtracts migration cost (bold solid). The wage profile in autarky is given by the thin solid line for comparison. Aggregate welfare is measured by the area under the autarky wage profile and open economy wage profile, respectively.

The main welfare implications of high-skilled migration are summarised as follows:

Proposition 5.3.1 *International migration leads to aggregate production gains, and to losses in aggregate welfare. Furthermore, all individuals are worse off in the laissez-faire migration equilibrium than in the autarky equilibrium.*

Following the approach of Benhabib and Jovanovic (2012) we now look at the social planner equilibrium. The social planner can freely choose the migration cutoff $\tilde{s}$ taking as given migration cost c, but disregarding individuals' migration incentives, which link $\tilde{s}^{\text{lf}}$ to c in the *laissez-faire* equilibrium. Maximising Eq. (5.5) with respect to $\tilde{s}$ gives the optimal migration cutoff $\tilde{s}^{\text{sp}}$ and hence the socially optimal extent of migration as a function of c:

$$\tilde{s}^{\text{sp}} = \frac{1}{2} + \frac{2c}{A}. \tag{5.6}$$

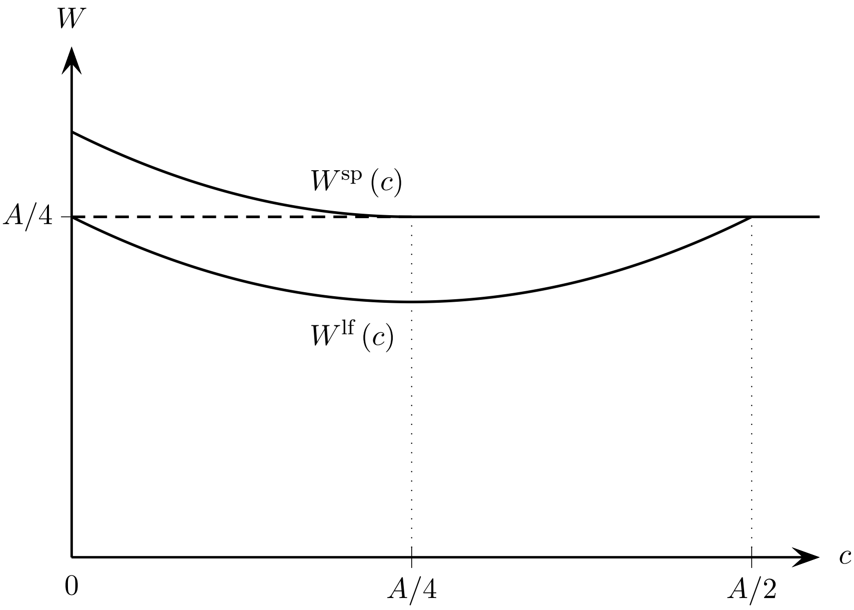
Hence, while there is "too much" migration in the laissez-faire equilibrium due to the negative migration externality, the optimal level of migration is strictly positive if migration costs are sufficiently low. Note also that $\tilde{s}^{\text{sp}} > 1/2$ and therefore it is never socially optimal to have more than half the population emigrating. It furthermore follows from Eq. (5.6) that zero migration is enforced by the social planner ($\tilde{s}^{\text{sp}} = 1$) whenever $c \geq A/4$.

We can compare welfare in the laissez-faire and social planner scenarios by substituting the respective migration cutoffs from Eqs. (5.4) and (5.6) into Eq. (5.5), thereby expressing aggregate welfare in each scenario as a function of migration cost:

$$W^{\text{lf}}(c) = \frac{A}{4} - c\left(\frac{1}{2} - \frac{c}{A}\right),$$
$$W^{\text{sp}}(c) = \frac{5A}{16} - c\left(\frac{1}{2} - \frac{c}{A}\right),$$

and it is easily checked that $W^{\text{sp}}(c)$ is strictly larger than autarky welfare $A/4$ for all non-prohibitive levels of c. The relationship between aggregate welfare and migration cost in the laissez-faire equilibrium and the social-planner equilibrium is illustrated in Figure 5.3.

Figure 5.3: *Aggregate welfare*

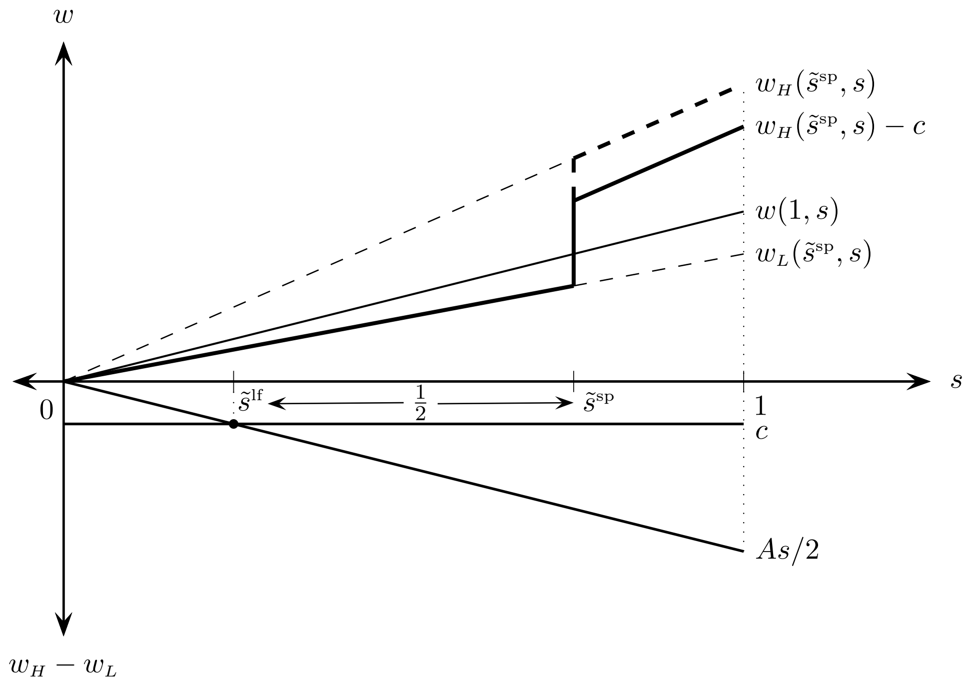


We now look at the effect that a socially optimal level of international migration has on individual wages. Clearly, non-migrants are worse off with any level of high-skill emigration, since the expected quality of their co-workers falls. Hence, we can restrict our attention to comparing the expected net wage of migrants in the social optimum with the respective wage in autarky. The net wage of migrants is given by

$$w_H(\tilde{s}^{\text{sp}}, s) - c = \frac{A(1+\tilde{s}^{\text{sp}})s}{2} - c,$$

and, substituting for $\tilde{s}^{\text{sp}}$, it is immediate that there is a wage gain relative to autarky for migrants with skill, $s > 4c/(4c + A)$. Simple algebra shows that this threshold value is strictly smaller than $\tilde{s}^{\text{sp}}$ as derived in Eq. (5.6), and therefore in the social optimum all migrants are better off than under autarky. Figure 5.4, which is directly analogous to Figure 5.2 (but for expositional purposes considers a smaller migration cost c) illustrates this. In constructing Figure 5.4, we use the fact that from our results in Eqs. (5.4) and (5.6) we know that $\tilde{s}^{\text{sp}} = \tilde{s}^{\text{lf}} + 1/2$. Furthermore, the size of the jump in the wage profile in the upper quadrant at $\tilde{s}^{\text{sp}}$ is determined by the the wage gain for the marginal migrant, which is determined in the lower quadrant. Proposition 5.3.2 summarises the results:

Figure 5.4: *Social planner equilibrium*



Proposition 5.3.2 *The socially optimal level of migration is strictly lower than in the laissez-faire equilibrium, if the latter features positive migration levels. For* $c < A/4$ *the socially optimal level of migration is strictly positive. In the social optimum, all migrants are better off than under autarky, while all non-migrants are worse off.*

The social optimum can alternatively be implemented by a tax on migration by both countries. In this case, individual incentives to migrate are again relevant, of course. We assume that a country's tax revenue is distributed equally to all nationals, independent of their residence, and hence does not affect the migration decision. Note that what countries care about in our setup is emigration, not immigration: Immigrants do not interact with natives, and hence have no effect on their wage rate, while emigration reduces the quality of matches available for those left behind. Hence, a government in our framework would set an emigration tax, not an immigration tax. Condition (5.4) now holds in a modified form, with effective (tax-inclusive) migration cost $c + t$ replacing c:

$$\tilde{s} = \frac{2(c+t)}{A}.$$

Substituting for $\tilde{s}$, using $\tilde{s}^{\text{sp}}$ from Eq. (5.6), implies $t^{\text{sp}} = A/4$. Notably, the optimal emigration tax rate does not depend on whether it is set cooperatively between countries, or non-cooperatively. This is due to the fact, mentioned above, that a country's welfare is independent of the extent of immigration (which is the only variable affected by the other country's emigration tax).

Figure 5.5: *Equilibrium with optimal emigration tax*

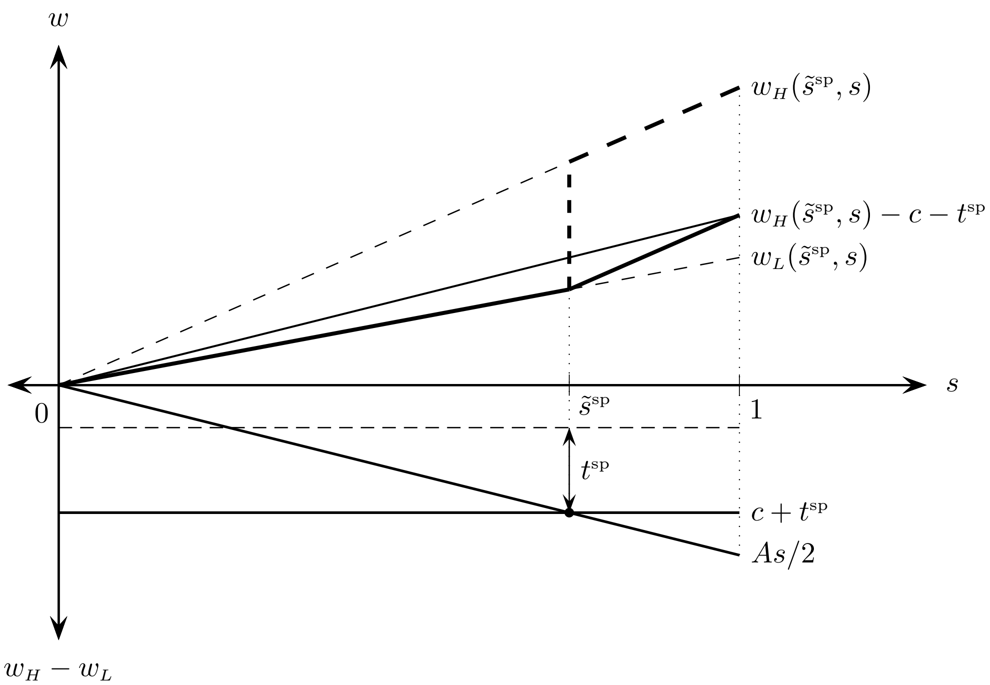


Figure 5.5 shows the resulting distribution of wages, where as before the bold dashed line gives the distribution of gross wages, and the bold solid line gives the distribution of net wages, subtracting effective migration cost $c + t^{\text{sp}}$. While in principle Figure 5.5 resembles Figure 5.2 from the laissez-faire equilibrium, with $c + t^{\text{sp}}$ substituted for c, there is one crucial difference: The migration equilibrium now yields tax revenue, which is equally redistributed among natives. The resulting transfer-inclusive wage is not shown in Figure 5.5 in order to avoid clutter, but it is clear that the transfer leads to a parallel upward shift in the net-wage profile. Consequently, individuals with the highest abilities and individuals with the lowest abilities are better off than in autarky: For both groups the absolute pre-transfer losses relative to autarky are small, as shown above, and therefore their transfer-inclusive incomes are higher than in autarky. It can be

shown analytically that this simple tax-transfer scheme does not make everyone better off than in autarky, and hence individuals with intermediate abilities (the most high-skilled non-migrants and the least skilled migrants) see their transfer-inclusive net wages fall.[11]

5.4 Extensions

Five important assumptions of the model presented in Sections 5.2 and 5.3 are that *(i)* individual ability of all workers is unobservable, *(ii)* migrants' skills are perfectly transferable across countries, *(iii)* migration is the only available signalling device, *(iv)* internationally mobile labour is the only factor of production, and *(v)* countries are ex ante identical in all respects. We now consider extensions of our model, where we relax these five assumptions one at a time. In doing so, we focus on the most interesting implications of the respective extension. In Section 5.4.1 we consider the case where the ability of individuals becomes observable with a positive probability. In Section 5.4.2 we allow for imperfect transferability of migrants' skills. In Section 5.4.3 we introduce education as an alternative signalling device. In Section 5.4.4 we add an internationally immobile factor of production to the model. In Section 5.4.5 we consider country asymmetries.

5.4.1 Imperfect observability of skill

As discussed earlier, one key stylised fact that our benchmark model does not capture well is the imperfect segregation between high-skilled migrants and non-migrants in the workplace, as documented by Hellerstein and Neumark (2008), Andersson, García-Pérez, Haltiwanger, McCue, and Sanders (2010), Aslund and Skans (2010). In our benchmark model the probability of a given migrant being matched with another migrant is equal to one, while the empirical studies find matching rates in excess of those that would be found under random matching, but significantly smaller than one. We now demonstrate that imperfect observability of skill leads to exactly the same outcome in our model.[12]

[11] The proof is delegated to Appendix A.21.

[12] Hendricks (2001) introduces the possibility of cross-matching between migrants and natives by assuming that an exogenous fraction of migrants is indistinguishable from natives.

For the sake of continued tractability we model the imperfect observability of abilities in a parsimonious and stylised way. Consider the following sequence of events. Before individuals decide about migration their abilities are revealed with probability $p \in (0, 1)$. Then, as in our baseline model, individuals decide whether to migrate, incurring migration cost $c > 0$, or to stay put. This decision is based on a comparison of expected incomes. Once migration has taken place, with probability $q \in (0, 1)$ the abilities of those whose skills have been private knowledge so far, are revealed. Finally, firms hire workers and production takes place.

Before considering a worker's migration decision in this changed environment, we have to derive the wage schedule for workers with observable skills. The firm's profit maximisation problem can analogously to Eq. (5.2) be written as

$$\max_{s_1, s_2} \pi (s_1, s_2) = 2As_1 s_2 - w (s_1) - w (s_2), \tag{5.7}$$

in which s_l, $l = 1, 2$, refers to the skill of a worker performing task $l = 1, 2$, while $w (s_l)$ denotes the wage paid to this workers. The solution to the profit maximisation problem is given by the following lemma:

Lemma 5.4.1 *If workers' skills are perfectly observable, firms maximise their profits by hiring only workers with exactly the same skill level.*

Proof Positive assortative matching of workers within firms follows immediately from the supermodularity of Eq. (5.1), see Kremer (1993).

Using the zero profit condition as well as the result on positive assortative matching in Eq. (5.7), the wage rate of a worker with observable skill level s is given by

$$w (s) = As^2. \tag{5.8}$$

Now it is easy to see that individuals with ex ante observable skills have no incentive to migrate, irrespective of their skill level: They are positively assortatively matched in any case, leaving them with a wage rate as given by Eq. (5.8), and by staying put they can save migration cost c. For workers whose skill is unobservable ex ante, an analogous logic to Section 5.2 applies: They know that with probability $1 - q$ their skill level remains unobservable ex post, in which case a switch from low-skill group L, with $\bar{s}_L = \tilde{s}/2$, to high-skill group H, with $\bar{s}_H = (1+\tilde{s})/2$, yields a wage gain of $As/2$. However, with probability q their skill level is revealed ex post and the worker earns the same wage at home and abroad. Hence, the expected wage gain of switching from group L to H amounts to $(1-q)As/2$. For the indifferent worker with skill $\tilde{s}$ condition $(1-q)A\tilde{s}/2 = c$ must hold, giving a migration cutoff

$$\tilde{s}^{\mathrm{lf}} = \frac{2c}{(1-q)A}.$$

Comparison to Eq. (5.4) from the benchmark model shows that a positive probability q of a migrant's skill being revealed ex post increases the migration cutoff, i.e. reduces the incidence of migration among those with ex ante undisclosed skill levels.

We now illustrate the degree of workplace segregation predicted by our model. Consider first the probability that a randomly picked migrant would have another migrant as a co-worker under random matching. This would happen with a probability equal to the migrants' population share, which is $(1-p)(1-\tilde{s}^{\mathrm{lf}})$. Now consider the same probability predicted by the model. With probability $(1-q)$ the migrant's skill is private knowledge, in which case he is matched with another migrant with probability one. With probability q his skill is revealed ex post, and he is matched with a co-worker of identical skill. Within the relevant group of individuals whose skill has been revealed, the share of migrants is $(1-p)q/[(1-p)q+p]$, where $(1-p)q$ is the share of migrants of known skill in the overall population at this skill level, and p is the share of natives in the overall population at this skill level.

Hence, in our extended model the probability for a random migrant to be matched with another migrant is equal to

$$\mathrm{Prob}(p,q) = 1 - q + q\left[\frac{(1-p)q}{(1-p)q+p}\right],$$

and it is easily shown that $\mathrm{Prob}(p,0) = 1$, $\mathrm{Prob}(p,1) = 1-p$, and $\partial\mathrm{Prob}/\partial q < 0$. Hence, the probability for a random migrant to be matched with another migrant is higher than under random matching. Interestingly, for a given migrant the probability of being matched with another migrant does not depend on his skill level s. This is also compatible with the results from Hellerstein and Neumark (2008), Andersson, García-Pérez, Haltiwanger, McCue, and Sanders

(2010), and Aslund and Skans (2010), who find that workplace segregation is at most weakly related to skill levels. Summing up, we have the following result:

Proposition 5.4.2 *The probability of migrants to have a co-worker who is also a migrant does not depend on their skill level, and it is furthermore smaller than one, but larger than under random assignment of workers into workplaces.*

5.4.2 Imperfect transferability of skills

As discussed in Section 5.2, our baseline model implies zero overlap in the skill range of migrants and natives: Migrants are always at the top of the destination country's wage distribution, while natives are at the bottom. This outcome is a consequence of the assumption in the benchmark version of our model that skills are perfectly transferable between countries. In accordance with results from Mattoo, Neagu, and Özden (2008) and Chiswick and Miller (2009), who show that immigrants in the US are more likely to suffer from occupational "underplacement" than natives, we now allow for a less than perfect transferability of workers' pre-migration skills. In particular, we assume that migrants can transfer only a fraction θ of their skills, while the fraction $1-\theta$ of skills is country specific and therefore becomes obsolete when going abroad. The migration arbitrage condition then reads

$$\frac{A\tilde{s}}{2}\left[(1+\tilde{s})\theta^2-\tilde{s}\right]=c,$$

where we have substituted $\bar{s}_L=\tilde{s}/2$ and $\bar{s}_H=\theta\,(1+\tilde{s})\,/2$. Solving for the *laissez-faire* migration cutoff $\tilde{s}^{\mathrm{lf}}$ yields

$$\tilde{s}^{\mathrm{lf}}=\frac{A\theta^2-\sqrt{A^2\theta^4-8\,(1-\theta^2)\,c}}{2A\,(1-\theta^2)}, \tag{5.9}$$

where $\tilde{s}^{\mathrm{lf}}\in(0,1)\ \forall\ c\in\left(0,A\left(2\theta^2-1\right)/2\right)$, which implies that an economically meaningful solution requires $\theta\in(\sqrt{1/2},1]$. Differentiating Eq. (5.9), we find that given our parameter constraint for θ, we have $\partial\tilde{s}^{\mathrm{lf}}/\partial\theta<0$. Thus, as one would reasonably expect, lower skill transferability θ weakens the incentive to migrate. There is now an overlap of the skill distributions by migrants and natives, respectively, since the lowest-skill immigrant has skill level $\theta\tilde{s}^{\mathrm{lf}}$, while the highest-skill native has skill level $\tilde{s}^{\mathrm{lf}}$. Proposition 5.4.3 sums up.

Proposition 5.4.3 *If skills are imperfectly transferable internationally, the skill distributions of migrants and natives overlap.*

5.4.3 Migration vs. education as signalling devices

While in our baseline model the only way for individuals to signal their true skill is by costly migration, in reality there is of course a wide range of possible signals, with education being probably the best known example, as already outlined by Spence (1973). We now analyse whether the presence of costly education as an alternative signalling device limits the importance of our signalling story in explaining the phenomenon of two-way migration.

Similar to migration, education involves a fixed cost, $c_e > 0$, and workers can now choose whether to emigrate, to get an education, or to do neither. Focussing on the signalling aspect of education, it is assumed that education does not alter workers' skills. Firms observe both signals and use this information to form more efficient matches at the workplace. The equilibrium is derived in the same way as in the baseline model. The results are summarised in the following proposition.

Proposition 5.4.4 *With costly migration and costly education as two alternative ways for workers to signal their skill, and provided the cost of neither signal is prohibitive,*

(i) high-skilled workers select into the costlier signal, while medium-skilled workers select into the less costly signal, whenever costs for the two signals are sufficiently different,

(ii) high-skilled workers select into the costlier signal, and the other signal is not chosen, whenever costs for the two signals are sufficiently similar.

Proof See Appendix A.22.

The intuition for Proposition 5.4.4 is the following. If the costs of the two signals are sufficiently different, high-skilled workers use the costlier signal to get separated from co-workers with lower skills. Medium-skilled workers are deterred from the costlier signal, but they have an incentive to get separated from low-skilled workers, which is achieved by selecting into the signal with the lower cost. Now consider the case where the costs of the two signals become more similar, by holding the cost of the cheaper signal constant, while the cost of the more expensive signal gradually declines. As the costlier signal is easier to afford, the group of individuals choosing the cheaper signal shrinks at both ends: The most high-skilled in this group now select the expensive signal. This in turn makes it less attractive for everybody else to be in this group, causing workers to drop out at the lower end as well. With converging costs of the two signals, this mechanism eventually leads to the disappearance of the group choosing the cheaper signal.

Figure 5.6: *Possible equilibria with two alternative signals*

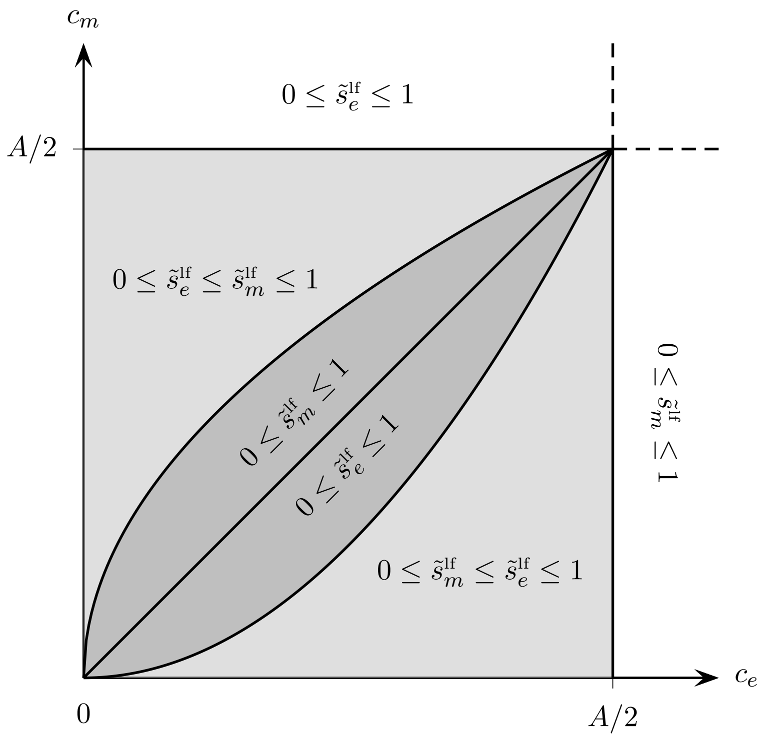


Figure 5.6 illustrates this result. The two curves enclosing the dark lens are given by $c_m = 2c_e^2/A$ and $c_e = 2c_m^2/A$, respectively. All parameter constellations within this lens represent cases in which the costs for the two signals are similar, and in these cases only the costlier signal is used. For combinations of c_m and c_e outside the lens, but inside the light, grey square both signals coexist with the high-skilled (medium-skilled) workers using the expensive (cheap) signal. If one of the signals is prohibitively expensive, i.e. $c_m \geq A/2$ or $c_e \geq A/2$, only the cheaper one is used. If both are too costly, none is used.

To sum up, in general, adding education as an alternative signalling device does not rule out the use of migration as a signal. In fact for the largest part of the relevant parameter space both signals coexist. In particular it is shown in the appendix that, if $0 < c_e < c_m < A/2$, the resulting migration cutoff $\tilde{s}_m$ is the same as in Eq. (5.4). Only for parameter combinations leading to $0 < 2c_e^2/A < c_m < c_e < A/2$ education completely replaces migration as a signalling device.[13]

5.4.4 Internationally immobile factors of production

In this subsection, we add internationally immobile capital to our model. Capital is modelled as an essential input in all firms, and hence we introduce an interaction between migrants and domestic factors of production that is standard in most migration models (cf. Berry and Soligo, 1969; Borjas, 1999), but has not been a feature of our basic model. The production technology is unchanged with respect to labour, i.e. there are two tasks, which have to be performed by exactly one worker each, and following Kremer (1993) we assume that capital is combined with labour in a Cobb-Douglas fashion. The resulting production function is given by

$$y = f(s_1, s_2, k) = 2As_1s_2k^{\alpha}, \tag{5.10}$$

with $\alpha \in [0, 1]$ denoting the partial production elasticity of capital and k being the per capita capital stock used in production. With firms knowing only the average skill within the groups, L and H, Lemma 5.2.1 implies positive assortative matching of group members. The profit maximising level of capital depends on whether the firm employs individuals from group H or L, and we show in the appendix that the amount of capital used by either type of firm is given by:

$$k_L = \left[\tilde{s} + (1-\tilde{s})\left(\frac{1+\tilde{s}}{\tilde{s}}\right)^{\frac{2}{1-\alpha}}\right]^{-1} \bar{k}, \tag{5.11}$$

$$k_H = \left[(1-\tilde{s}) + \tilde{s}\left(\frac{\tilde{s}}{1+\tilde{s}}\right)^{\frac{2}{1-\alpha}}\right]^{-1} \bar{k}, \tag{5.12}$$

[13] If the cost of education declines in a worker's skill, such that the effective cost of eduction for an individual with skill s equals c_e/s instead of c_e, it can be shown that the equilibrium is of the type $0 < \tilde{s}_m^{\text{lf}} < \tilde{s}_e^{\text{lf}} < 1$ for all $0 < c_m < c_e < A/2$.

where $\bar{k}$ is the average capital stock in the economy. It is easily checked that $k_H \geq \bar{k} \geq k_L$. Hence, firms employing workers of a higher expected ability, which in equilibrium will be firms employing migrants, have a higher capital intensity.

In analogy to Section 5.2, wages are determined by splitting available revenue (now the difference between total firm revenue and payments to capital) equally between the two workers. Capital returns are distributed equally among the nationals of a country, and hence capital ownership does not distort the decision to migrate. In analogy to the baseline model, the laissez-faire migration equilibrium is then determined by the condition that the wage gain for the marginal migrant is equal to the migration cost. We get

$$\tilde{s}^{\mathrm{lf}} = \frac{2c}{A(1-\alpha)\bar{k}^{\alpha}} \left(\Phi\right)^{-1}, \tag{5.13}$$

with

$$\Phi \equiv (1+\tilde{s}^{\mathrm{lf}}) \left(\frac{k_H}{\bar{k}}\right)^{\alpha} - \tilde{s}^{\mathrm{lf}} \left(\frac{k_L}{\bar{k}}\right)^{\alpha} \geq 1,$$

where the inequality is strict whenever $\alpha > 0$. Comparison with Eq. (5.4) shows that the relative size of the laissez-faire migration cutoffs in the two models depends on two effects. A larger value for $(1-\alpha)\bar{k}^{\alpha}$ increases migration flows since the migration cost falls in relation to average income. The second effect is given by Φ^{-1}, and it shows that an additional incentive to migrate exists in the extended model, which stems from the reallocation of domestic capital towards firms employing (more productive) migrants.

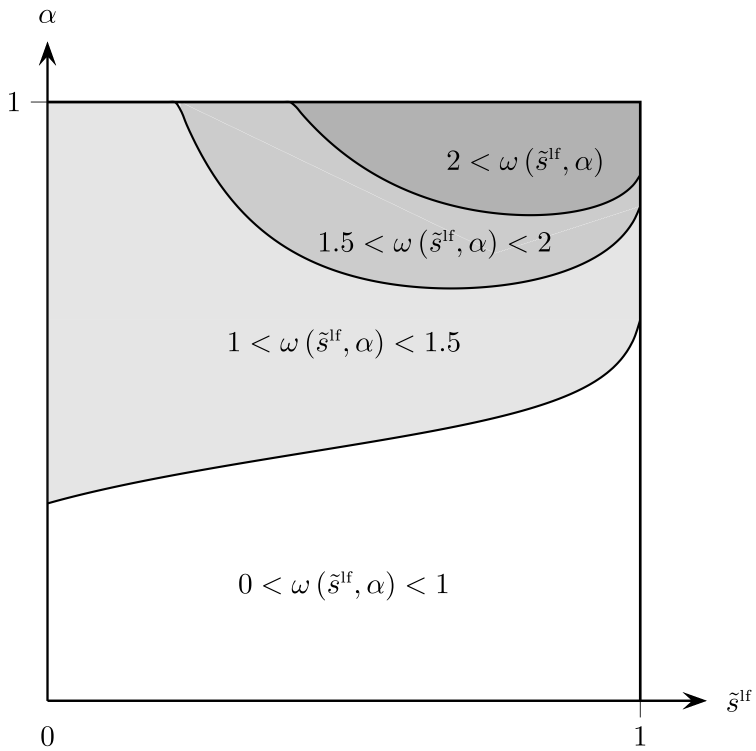
We now turn to the welfare implications that migration has in the framework with capital just described. Going through the same steps as in the baseline model, we find that aggregate welfare in the laissez-faire migration equilibrium is given by

$$W(\tilde{s}^{\mathrm{lf}}, \alpha) = \frac{A\left\{k_H^{\alpha} - \left[2\Phi(1-\alpha)\bar{k}^{\alpha} - k_H^{\alpha}\right]\tilde{s}^{\mathrm{lf}}(1-\tilde{s}^{\mathrm{lf}}) - (k_H^{\alpha} - k_L^{\alpha})\left(\tilde{s}^{\mathrm{lf}}\right)^3\right\}}{4},$$

and it is easily checked that autarky welfare is equal to $W(1,\alpha) = A\bar{k}^{\alpha}/4$. We can now compute the relative welfare levels in the migration equilibrium and in autarky, $\omega\left(\tilde{s}^{\mathrm{lf}}, \alpha\right) \equiv W\left(\tilde{s}^{\mathrm{lf}}, \alpha\right)/W\left(1, \alpha\right)$, where aggregate migration gains exist whenever $\omega(\tilde{s}^{\mathrm{lf}}, \alpha) > 1$.

Figure 5.7 provides a contour plot of $\omega\left(\tilde{s}^{\mathrm{lf}}, \alpha\right)$ for all combinations of $\tilde{s}^{\mathrm{lf}}$ and α, where combinations that lead to $\omega\left(\tilde{s}^{\mathrm{lf}}, \alpha\right) > 1$ are highlighted in different shades of grey. All other

Figure 5.7: *Aggregate welfare in a model with capital*



combinations lead to aggregate welfare losses from migration. We find that in contrast to our baseline model that abstracts from complementarities in production between internationally mobile and immobile factors, there exists now a non-trivial parameter space where welfare losses from the negative migration externality are overcompensated by the efficiency gains resulting from the reallocation of capital towards migrant-employing firms. The results are summarised as follows:

Proposition 5.4.5 *For high (low) values of α the model features aggregate welfare gains (losses) from international migration.*

Turning to the social planner's solution, one can show that the socially optimal level of migration will be lower than the one in the *laissez-faire* equilibrium given by Eq. (5.13).[14] It is easy to

[14]The proof is deferred to Appendix A.24.

see why: Adding capital to the model opens up a new channel for gains from migration, but does not add a new distortion. Hence, the migration externality remains the only distortion in the model. As an immediate consequence migration levels in the *laissez-faire* equilibrium will in general be too high.

5.4.5 Country asymmetries

We now extend our baseline model by assuming $A_D \neq A_F$, where A_D denotes the technology level of the domestic economy while A_F refers to the corresponding technology parameter in the foreign economy. Recalling Eq. (5.3), the two country-specific indifference conditions for the marginal migrant can be written as

$$\frac{A_i \tilde{s}_i}{2}\left[\frac{A_j}{A_i}(1+\tilde{s}_i) - \tilde{s}_i\right] = c \quad \forall \quad i,j \in \{D,F\} \quad \text{with} \quad i \neq j, \tag{5.14}$$

where we have used $\bar{s}_{Li} = \tilde{s}_i/2$ and $\bar{s}_{Hi} = (1+\tilde{s}_i)/2$. Solving for $\tilde{s}_i^{\text{lf}}$ yields

$$\tilde{s}_i^{\text{lf}} = \frac{A_j - \sqrt{A_j^2 + 8\,(A_j - A_i)\,c}}{2\,(A_i - A_j)} \quad \forall \quad i,j \in \{D,F\} \quad \text{with} \quad i \neq j. \tag{5.15}$$

It is now easy to check that the technologically superior country experiences net immigration, i.e. for $A_j > A_i$ we have $\tilde{s}_j^{\text{lf}} > \tilde{s}_i^{\text{lf}}$. Moreover, it follows from differentiating Eq. (5.14) that $\partial\tilde{s}_i^{\text{lf}}/\partial A_i > 0 > \partial\tilde{s}_i^{\text{lf}}/\partial A_j$ if countries are not too dissimilar, i.e. if $2/3 < A_D/A_F < 3/2$. This is the case we focus on henceforth. Thus, emigration increases if the technology in the destination country gets better, while it falls if the same occurs in the source country. The prohibitive level of migration cost is now also country-specific: Setting $\tilde{s}_i^{\text{lf}} = 1$ in (5.15), we find that emigration occurs from country i whenever $c < (2A_j - A_i)/2$.

Turning to the welfare implications of migration, aggregate welfare of nationals from country $i \in \{D,F\}$ can be expressed analogously to Eq. (5.5) as

$$W_i(\tilde{s}_i, c) = \frac{(A_i - A_j)(\tilde{s}_i)^3}{4} + \frac{A_j[1+\tilde{s}_i(1-\tilde{s}_i)]}{4} - (1-\tilde{s}_i)c,$$

for all $i,j \in \{D,F\}$ with $i \neq j$. In analogy to the baseline model we can use the link between migration cost and the laissez-faire migration cutoff in Eq. (5.15) to express aggregate welfare as a function of $\tilde{s}_i^{\text{lf}}$ alone:

$$W_i^{\text{lf}}(\tilde{s}_i^{\text{lf}}) = \frac{A_i(\tilde{s}_i^{\text{lf}})^2(2-\tilde{s}_i^{\text{lf}})}{4} + \frac{A_j[1-(\tilde{s}_i^{\text{lf}})^2](1-\tilde{s}_i^{\text{lf}})}{4}. \tag{5.16}$$

Migration leads to aggregate welfare gains for the nationals of country i, whenever $W_i(\tilde{s}_i^{\mathrm{lf}}) > W_i(1) = A_i/4$ for $\tilde{s}_i^{\mathrm{lf}} \in (0,1)$, where (5.15) can be used to derive the necessary condition on the cost of migration. We find the following:

Proposition 5.4.6 *Aggregate welfare is lower in a migration equilibrium than under autarky for nationals of the country with the better technology. For nationals of the technologically inferior country, aggregate welfare gains from migration exist if migration costs are sufficiently low.*

Proof See Appendix A.25.

Relative to the baseline model, in which all individuals lose from trade in the laissez-faire equilibrium, country asymmetries result in an additional welfare effect that is positive for migrants from the technologically inferior country (since they use a more efficient technology in the destination country) and negative for migrants from the other country. It is therefore intuitively plausible that only nationals from the technologically inferior country may gain in the aggregate from migration.[15]

With the asymmetric version of our model at hand we can now return to Figure 5.1, which compares two-way migration within the EU15 and the OECD. Using Eq. (5.15) and focussing (without loss of generality) on the case $A_j \geq A_i$, it is now possible to compute the familiar index of bilateral balance in migration:

$$B_{ij}(\tilde{s}_i^{\mathrm{lf}}, \tilde{s}_j^{\mathrm{lf}}) = \frac{2\min\left(\mathrm{Em}_{ij}, \mathrm{Em}_{ji}\right)}{\mathrm{Em}_{ij} + \mathrm{Em}_{ji}} = \frac{2(1-\tilde{s}_j^{\mathrm{lf}})}{2-\tilde{s}_j^{\mathrm{lf}}-\tilde{s}_i^{\mathrm{lf}}}. \tag{5.17}$$

Note that, if countries are identical, i.e. $A_j = A_i = A$, we have $\tilde{s}_j^{\mathrm{lf}} = \tilde{s}_j^{\mathrm{lf}} = \tilde{s}^{\mathrm{lf}}$ and $B_{ij}(\tilde{s}_i^{\mathrm{lf}}, \tilde{s}_j^{\mathrm{lf}})$ in Eq. (5.17) takes a value of one. Moreover, it is straightforward to show that $B_{ij}(\tilde{s}_i^{\mathrm{lf}}, \tilde{s}_j^{\mathrm{lf}})$ declines monotonically as $A_j - A_i$ increases: As countries become more dissimilar migration becomes less balanced. This is in accordance with the results in Table 1, which show that migration of tertiary educated individuals between EU15 country pairs is more balanced than between country pairs in the more heterogeneous group of OECD countries.

[15] Notably, the negative migration externality discussed in Section 5.3 is also present here. In particular we can show that the migration cutoffs $\tilde{s}_i^{\mathrm{sp}}\ \forall\ i = D, F$ that an omniscient social planner would choose are strictly higher than the ones from Eq. (5.15). The mathematical proof is deferred to Appendix A.26.

5.5 Temporary migration

While so far the focus mainly has been on permanent migration, we now acknowledge the fact that international migration today is increasingly seen as a temporary phenomenon (Dustmann and Glitz, 2011).[16] The theoretical migration literature thereby distinguishes between two broad trends explaining the temporary character of international migration: Low-skilled guest-worker migration (cf. Ethier, 1985; Djajic and Milbourne, 1988; Djajic, 1989, 2010, 2013a,b; Dustmann and Kirchkamp, 2002; Mesnard, 2004) and high-skilled student migration (cf. Dustmann and Weiss, 2007; Dustmann, Fadlon, and Weiss, 2011; Dustmann and Glitz, 2011). Common to both strands of the literature is the analysis of temporary migration in an asymmetric two-country setup, in which individuals emigrate from a developing country to a more advanced economy, in order to invest either in physical or in human capital accumulation. Both types of investments thereby tend to pay off at higher rates in the developing economy, and it is this difference in returns, which rationalises the observed return migration towards the seemingly less attractive location.

While this combination of country-specific push and pull factors provides a convincing explanation for temporary migration between asymmetric countries, it seems less clear what drives temporary migration between rather similar countries. To answer this question we develop a theory of high-skilled temporary *and* permanent migration between two identical countries. Individual (return) migration decisions thereby are analysed in a framework with two overlapping generations of heterogeneous workers, which differ with respect to their unobservable skills.[17]

[16] Comparable data on bilateral temporary migration and in particular on return migration is scant. The OECD reports average re-emigration rates for a small set of selected European countries and the US which vary from 19.1% for the US up to 60.4% for Ireland (cf. Table III.1. in OECD, 2008b). More recently, Gibson and McKenzie (2012) surveyed 4,131 high-talented top-performers from five typical "brain drain" countries (Ghana, Micronesia, New Zealand, Papua New Guinea and Tonga), which graduated from high school between 1976 and 2004. In their sample, 65% of all respondents have ever migrated abroad, while 36% currently lived abroad when the survey was conducted. Case studies on the return migration from single countries exist among others for the US (cf. Borjas and Bratsberg, 1996), the UK (cf. Dustmann and Weiss, 2007), and Canada (Aydemir and Robinson, 2008). A more detailed review over the respective literature is given in Dustmann and Glitz (2011).

[17] Traditionally, temporary migration has either been studied in life-cycle models à la Djajic and Milbourne (1988), Djajic (1989, 2010), Dustmann and Kirchkamp (2002), and Mesnard (2004), or in models with overlapping

The production process requires the formation of teams as in Kremer (1993), and, hence, induces high-skilled workers to use costly one- or two-period stays abroad in order to escape from potentially "bad" matches with less skilled domestic co-workers. A costly stay abroad thereby acts as a signalling device (cf. Spence, 1973), and it is the repeated use of this signalling option, which shapes the strategic selection of workers into temporary and permanent migration. Depending on the costs of staying abroad two possible migration equilibria exist. If the costs are low, the symmetric equilibrium in both countries features temporary and permanent migration. Thereby, medium-skilled workers select into temporary migration, while high-skilled workers decide to migrate permanently abroad. On the contrary, if the costs of living abroad are high, we only observe temporary migration of the most high-skilled individuals.

In traditional asymmetric-country frameworks welfare gains for (temporary) migrants more or less automatically result as a consequence of workers' arbitrage between internationally unintegrated national markets.[18] Focusing on a setup with two identical countries, we would not expect these kind of welfare effects to matter, and indeed the welfare effects in our model contradict conventional wisdom in so far as all workers (including the migrants) tend to be worse off in an *laissez-faire* equilibrium with temporary (and permanent) migration than in an equilibrium without migration. Instrumental for the associated aggregate welfare loss is a negative migration externality, which leads to excessive temporary and permanent emigration in the presence of wasteful migration costs. As a consequence, aggregate production gains, which result from the more efficient matching of natives and migrants at the firm level, are eaten up by the costs of living abroad. Of course this does not mean that all migration, temporary or permanent, is socially harmful and, hence, must be prohibited. Employing an omniscient, global social planner we find that, if the costs of living abroad are not too high, the socially optimal equilibrium may feature temporary *and* permanent migration, both – of course – at a smaller scale than in the *laissez-faire* equilibrium. The global-social-planer solution thereby – as we show – can be implemented by a carefully chosen combination of emigration tax and return subsidy, which

generations (cf. Galor, 1986; Karayalcin, 1994).

[18] As an example for this kind of arbitrage Dustmann (2001) refers to the higher purchasing power of assets accumulated in the host country, when used for consumption back in the home country (e.g. after retirement).

both countries independent from each other consider as socially optimal.

We consider three separate extensions to our model. In particular we show that the availability of alternative signalling devices (e.g. education, see Spence (1973)) not necessarily leads to a complete crowding out of temporary *and* permanent migration as signalling devices. It rather seems to be the case that in particular high-skilled workers combine several signals (e.g. education and migration) to achieve a more accurate signal of their otherwise unobservable skills. We moreover highlight the strategic effect, that follows from the sequential structure of initial emigration and later return decisions, and show that initial emigration is strategically reduced if the marginal emigrant anticipates that the most high-skilled co-migrants permanently stay abroad. Finally, in the third extension we explore under which conditions our model not only features an equilibrium with temporary or temporary *and* permanent migration, but also an equilibrium in which *only* permanent migration results.

Our analysis is motivated by the distinct pattern of temporary migration between rather similar countries. While existing asymmetric-country models would predict unidirectional (return) migration flows, which should be reflected in bilateral migration stocks, that are considerably less balanced for short than for long durations of stay, we show that the observed pattern of temporary migration is characterised by a surprisingly stable balance in bilateral migration stocks over varying durations of stay. Due to lack of comparable data on bilateral temporary migration and in particular on return migration, we follow Dustmann and Glitz (2011) and use the Database on Immigrants in OECD Countries (DIOC) (cf. OECD, 2008c) to analyse the duration of stay in bilateral migration between 15 OECD countries for which this information is available around the year 2000.[19] Admittedly, the dissection of bilateral migration stocks by duration of stay yields a rather crude measure for temporary migration and the results must be interpreted with caution. As pointed out by Dustmann and Glitz (2011) high shares of short-term migrants could either result from actual short migration durations or from large number of recent arrivals (as for example in the case of Italy or Spain). Moreover we hold no information on which type of status change (e.g. return migration, onward migration or naturalisation)

[19] Notable exceptions are the studies by Gibson and McKenzie (2011, 2012); Dustmann and Glitz (2011), which focus on return migration to (from) a small set of sending (receiving) countries.

is responsible for the variation of bilateral migration stocks over different durations of stay.[20] Despite these limitations the aggregate figures in Table 5.2 point to a considerable amount of temporary migration in particular among the high-skilled: Only the half of all high-skilled migrants live for more than 20 years in their destination country and, hence, can be considered as permanent migrants. As an additional information in Table 5.2 we also report the share of bilateral migration stocks that can be characterised as two-way. The share is measured by the Index of Bilateral Balance in Migration (IBBM), which for each country pair (i, j) is given by $B_{ij} \equiv 2\min(\mathrm{Em}_{ij}, \mathrm{Em}_{ji})/(\mathrm{Em}_{ij} + \mathrm{Em}_{ji})$, with Em_{ij} as the stock of emigrants from country i residing in country j (cf. Biswas and McHardy, 2005).[21] The numbers in Table 5.2 are the average values of the index for a maximum of $15 \times 14 = 105$ OECD country pairs, in a given skill group and for varying durations of stay.

Three important insights can be gained from the Index of Bilateral Balance in Migration (IBBM). (i) Bilateral migration stocks are surprisingly balanced: Almost 50% of all high-skilled migration can be considered as two-way migration. (ii) The balance in bilateral migration stocks is stable over different durations of stay, i.e. the index values change only marginally if different durations of stay are considered.[22] And finally: (iii) Both trends seem to be most pronounced for the migration of high-skilled individuals.

To highlight these findings we take another, more disaggregated view on the data and plot in Figure 5.8 the bilateral migration stocks of high-skilled individuals, which stay more than 5 years (above the 45°-line), against the bilateral migration stocks of high-skilled individuals, which stay less than 5 years (below the 45°-line). Most of the observations cluster along the 45°-line, which indicates a considerable balance in bilateral migration for both durations of stay,

[20]One of the few studies that distinguishes between return and onward migration is Nekby (2006), which finds that the return rates among re-emigrants from Sweden vary between 90% for migrants from Nordic sending countries and 30% for migrants from African origin countries.

[21]The construction of the index is directly analogous to the well-known Grubel-Lloyd index (cf. Grubel and G.Lloyd, 1975) measuring intra-industry trade in differentiated products, i.e. two-way trade in goods within the same industry. See Brülhart (2009) for a recent application of the Grubel-Lloyd index as a measure of intra-industry trade.

[22]Migration that lasts for less than one year to a large extend is driven by seasonal workers and young working holidaymakers (cf. OECD, 2008c), which would explain the rather low index values in this category.

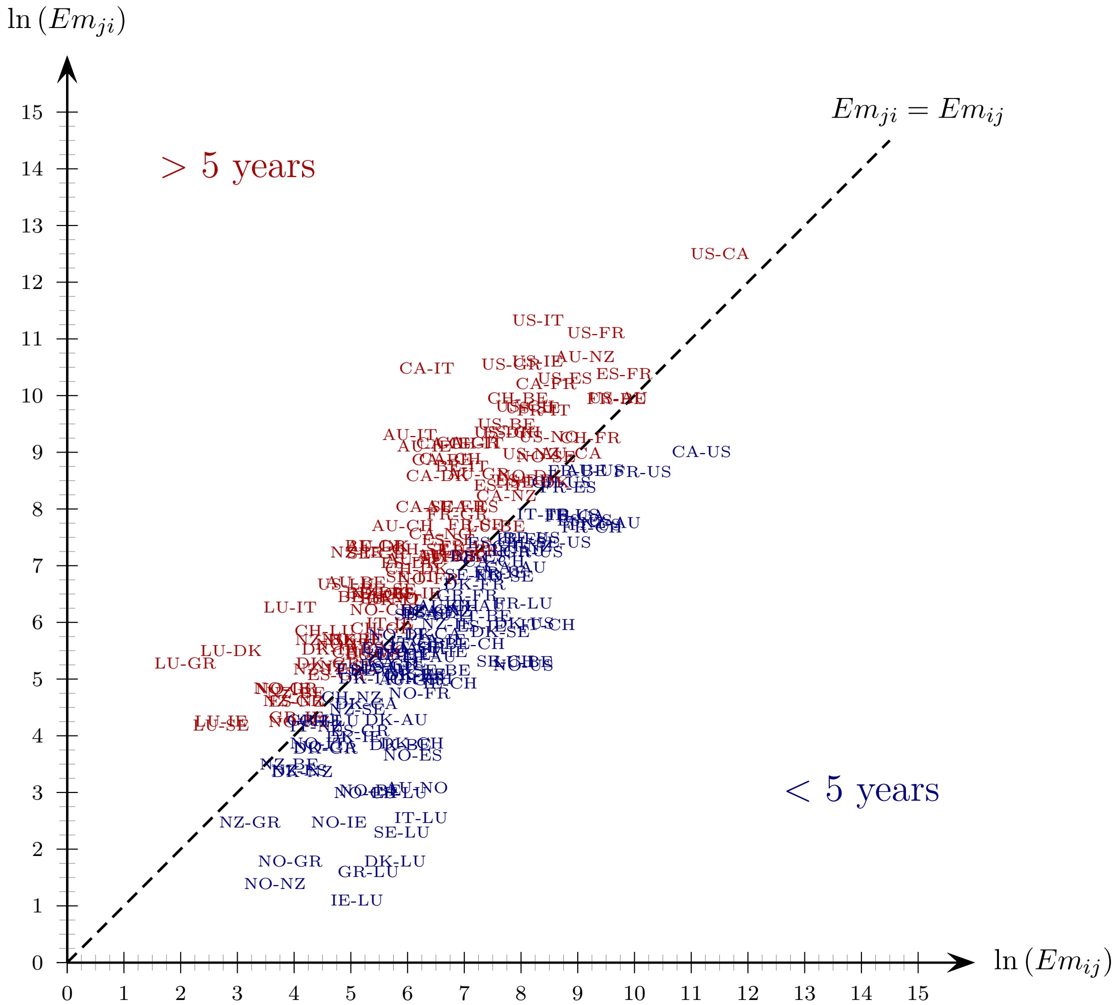
Table 5.2: *Migration shares and IBBMs for 15 OECD countries by duration of stay*

Duration of stay in years:	<1	1 - 3	3 - 5	5 - 10	10 - 20	>20	All
High skills:							
Share	0.07	0.09	0.06	0.11	0.16	0.51	1
IBBM	0.41	0.45	0.47	0.49	0.49	0.41	0.46
	(87)	(93)	(95)	(101)	(100)	(104)	(104)
Med. skills:							
Share	0.04	0.05	0.03	0.07	0.15	0.66	1
IBBM	0.36	0.41	0.42	0.44	0.43	0.25	0.33
	(80)	(91)	(96)	(99)	(100)	(101)	(105)
Low skills:							
Share	0.02	0.02	0.02	0.04	0.10	0.80	1
IBBM	0.30	0.39	0.42	0.46	0.40	0.18	0.24
	(59)	(78)	(77)	(86)	(93)	(94)	(102)
All skills:							
Share	0.04	0.05	0.04	0.07	0.13	0.67	1
IBBM	0.44	0.47	0.47	0.48	0.45	0.27	0.34
	(99)	(101)	(100)	(102)	(104)	(104)	(105)

Note: Table 5.2 reports the (average) shares of migrants by duration of stay along with the aggregate **I**ndex of **B**ilateral **B**alance in **M**igration (**IBBM**) for bilateral migration between 15 OECD countries (Australia, Belgium, Canada, Denmark, France, Greece, Ireland, Italy, Luxembourg, New Zealand, Norway, Spain, Sweden, Switzerland, and the United States) in 2000. Index values are computed separately at the country-pair level for varying durations of stay and different educational attainments. Computations thereby are based on a maximum of $(15 \times 14)/2 = 105$ country pairs, which we then aggregated up using the relative size of the respective bilateral migration stocks as weights. The number of country pairs available for the computation are reported in parenthesis below the respective index numbers.

and indeed the corresponding IBBMs take very similar values of 0.45 (0.46) for migrants that stay for more (less) than 5 years. Together these findings are noteworthy, given that traditional explanations based on country asymmetries in general would predict rather unbalanced bilateral migration stocks, which in particular for short durations of stay are considerably less balanced than for long durations of stay, (when temporary migration contribute to a lesser extent to the

Figure 5.8: *Two-way migration among 15 OECD countries by duration of stay*



Note: Figure 5.8 plots 101 out of $(15\times14)/2 = 105$ possible bilateral migration stocks for 15 OECD countries (Australia, Belgium, Canada, Denmark, France, Greece, Ireland, Italy, Luxembourg, New Zealand, Norway, Spain, Sweden, Switzerland, and the United States) in 2000, differentiated by duration of stay being either more or less than 5 years. Observations for LU-AU, LU-CA, LU-NO, and LU-NZ are missing. Note that for durations < 5 years (> 5 years) the net-emigration (net-immigration) country is named first. Hence the strict separation in above and below the 45Â°-line.

imbalance in bilateral migration stocks). Building up on this insight, in the following we develop a simple model of temporary *and* permanent migration between similar countries, which is able to explain the stable balance in bilateral migration stocks over different durations of stay as an outcome of strategic (return) migration decisions in a framework with asymmetric information. The stable balance of bilateral migration stocks thereby naturally results as workers in *both* countries make use of temporary and permanent stays abroad as signalling device, thereby generating balanced temporary *and* permanent migration flows. We thus complement the existing theoretical literature in providing a *novel* explanation for a symmetric selection into initial emigration *and* later return migration that is not based on country-specific push and pull factors.[23]

The idea to analyse migration in a framework with imperfect information is not new and has been used before to explain the differences in the performance of immigrants and natives (cf. Hendricks, 2001) and the geographic clustering of high-skilled workers (cf. Giannetti, 2001). However, to the best of our knowledge, we are the first to provide a simple framework, which not only allows us to analyse the *strategic* selection into initial emigration and later return migration, but also lends itself to a comprehensive welfare analysis and, hence, allows us to characterise optimal migration policies. Although the majority of models on temporary migration derive return decisions from country asymmetries (e.g. technology or endowment differences), there are a few exceptions, which emphasise the role of information asymmetries.[24] In an extension to his baseline migration model Hendricks (2001) analyses individual return decisions. Other than in our model migration decisions thereby depend on the cost of *moving* abroad instead of *living* abroad.[25] Although this distinction seems innocuous, it is crucial for the selection

[23]In Djajic and Milbourne (1988), Dustmann (2001) migrants earn a higher income abroad, but at the same time have a preference for consumption in their home country. In Djajic (1989), Dustmann (2001) a similar trade off results from the comparison of a higher income abroad versus a higher purchasing power at home. In Dustmann and Kirchkamp (2002), Mesnard (2004), and Djajic (2010) temporary migration results from a credit constraint and migrants trade off a higher labour income abroad against better investment possibilities at home. Finally, in Dustmann (2001), Dustmann and Weiss (2007), and Dustmann, Fadlon, and Weiss (2011) migrants have access to high-quality education abroad, which yields higher returns when supplied in the migrants' home country.

[24]For an analysis of temporary migration decisions in life-cycle models with uncertainty see Berninghaus and Seifert-Vogt (1988, 1993).

[25]Exploiting a rich survey on individual return migration to several typical "brain drain" countries, Gibson and McKenzie (2011, 2012) highlight the relative importance of the costs of *living* abroad, when compared to the cost of *moving* abroad.

into return migration. While in Hendricks (2001) the costly act of returning home generates a signal, in our framework it is the costly option to stay another period abroad which acts as a signalling device. Unlike Hendricks (2001), we thus find that only the most high-skilled initial emigrants decide to stay for another period abroad, while the least-skilled co-migrants return home, thereby accentuating the initial selection of the most high-skilled workers into emigration. A similar pattern is found by Stark (1995), who analyses return migration in an asymmetric country setting, in which workers' skills are more difficult to observe for firms in the destination country than for firms in the origin country (see also Katz and Stark, 1987). As in our model return incentives thereby arise in particular for low-skilled migrants, who no longer can expect to benefit from a pooling with high-skilled co-migrants once the information asymmetry between workers and firms for some reason is lifted. We add to this literature by endogenising the (partial) reinstatement of information symmetry as a consequence of workers' strategic migration/signalling decisions in a setting with two identical countries. Building up on this richer modelling environment we then, similar to Benhabib and Jovanovic (2012), ask for the globally optimal degree of (temporary and permanent) international migration. As a central difference to Benhabib and Jovanovic (2012), who focus on a classical efficiency-versus-distribution trade off, resulting from a *positive* human-capital externality as in Lucas (1988), the social planner in our model corrects for a *negative* migration externality, which in the *laissez-faire* equilibrium results in excessive emigration and aggregate welfare losses. Unlike in Benhabib and Jovanovic (2012) the social-planner equilibrium therefore features reduced, but non-zero levels of temporary and permanent migration. Responsible for this outcome are the aggregate production gains, which are generated by temporary and permanent migration through an improved matching of workers in the labour market.[26] The idea that temporary migration may contribute to a better matching of workers within firms by alleviating existing information asymmetries thereby naturally complements the existing literature (cf. Peri and Sparber, 2009; Peri, 2012), which focuses on the specialisation of natives and migrants on different sets of tasks (e.g. manual-physical tasks

[26]Empirical support for firm-level productivity effects of international migration comes from Trax, Brunow, and Suedekum (2012), who use administrative data from Germany to show that firms which are located in regional labour markets that are characterised by a high-degree of cultural diversity tend to be more productive.

vs. communication-language tasks) according to observable characteristics. Finally, we also show that the social-planner equilibrium can be implemented by a carefully chosen combination of emigration tax and return subsidy, which links our work to a recent and growing literature analysing optimal *temporary*-migration policies in a context of two asymmetric countries. (cf. Djajic and Michael, 2013; Djajic, Michael, and Vinogradova, 2012; Djajic, 2013b).

5.6 A simple model of strategic migration

Imagine two symmetric countries, producing a homogeneous *numéraire* good y, which is non-storable and can be traded costlessly at a world-market price normalised to $p \stackrel{!}{=} 1$.[27] In each period both countries are populated by two overlapping generations of workers, whose age we denote by $t = 1, 2$. Generation size is constant over time and, hence, can be normalised to unity without loss of generality. Workers in each generation are risk neutral and derive periodical utility from the consumption x of the *numéraire* good according to a linear periodical utility function $u(x) = x$. Lifetime utility then follows as the the sum over workers' periodical utilities at age $t = 1, 2$, given that workers are assumed to have no time preferences. Workers differ with respect to their skills, which follow from a uniform distribution over the interval $s \in [0, 1]$ and are assumed to be private information.

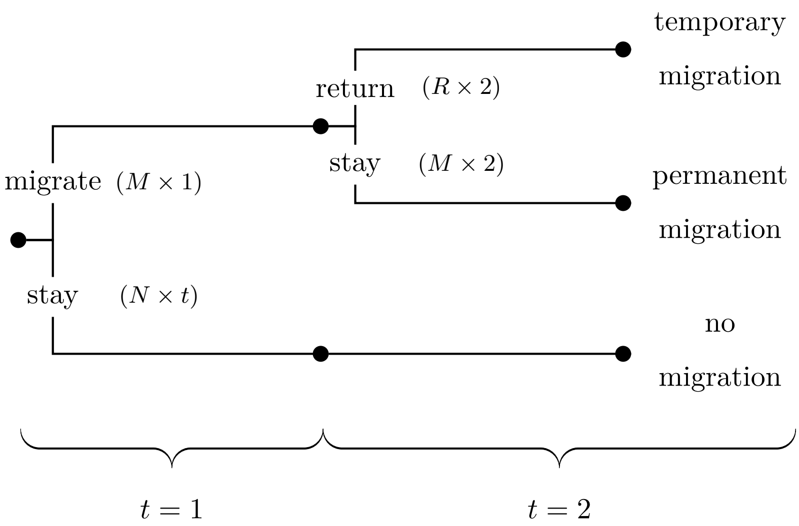
A worker's (return) migration decision can be sketched out as follows: At age $t = 1$, when still being young, the worker either stays put and gets employed at home or emigrates abroad and finds employment there.[28] Subsequently, a worker, who went abroad at age $t = 1$, at age $t = 2$ then either returns home or stays abroad for another (final) period. Figure 5.9 summarises the available migration options and identifies the resulting individual migration patterns.

Production takes place under perfect competition, using an "O-ring" production technology (cf. Kremer, 1993), which requires the processing of two tasks $l = 1, 2$, each to be performed by

[27] Given that countries are symmetric we suppress all country indices henceforth.

[28] For the sake of simplicity and consistent with typical life-cycle patterns of migration (cf. Sjaastad, 1962; Gallaway, 1969; Schwartz, 1976; Goss and Paul, 1986; Johnson, Voss, Hammer, Fuguitt, and McNiven, 2005) we do not consider late emigration at age $t = 2$ when workers are old.

Figure 5.9: *Migration and return decisions*



a single worker. Firm-level output follows correspondingly as:

$$y = f(s_1, s_2) = 2As_1s_2, \tag{5.18}$$

with $A > 0$ being a technology parameter and s_l denoting the skill level of the worker performing task $l = 1, 2$. Crucially, we have $\partial f(s_1, s_2)/\partial s_l > 0$ and $\partial^2 f(s_1, s_2)/\partial s_l s_{\hat{l}} > 0 \ \forall \ l, \hat{l} = 1, 2$ with $l \neq \hat{l}$, such that Eq. (5.18) is supermodular and workers enter production as complements.

In an equilibrium that features either temporary or permanent migration (cf. Figure 5.9) firms can identify individual workers as members of either the group of non-migrants N (at age $t = 1, 2$), the group of migrants M (at age $t = 1, 2$) or the group of returnees R (at age $t = 2$). This is the only information firms can base their hiring decision on, and this information is valuable since, as we show below, the average skills within the various subgroups (non-migrants, migrants, and returnees) are different. Taking into account these differences, firms then maximise their expected profits by choosing the optimal skill mix of their employees:

$$\max_{\bar{s}_1, \bar{s}_2} \ \pi\left(\bar{s}_1, \bar{s}_2\right) = 2A\bar{s}_1\bar{s}_2 - w\left(\bar{s}_1\right) - w\left(\bar{s}_2\right), \tag{5.19}$$

with $\bar{s}_l$, $l = 1, 2$ referring to the average skill of the group from which the worker performing task l is recruited, and $w(\bar{s}_l)$ being the wage paid to this worker. The solution to the profit maximisation problem in Eq. (5.19) is given by Lemma 5.6.1.

Lemma 5.6.1 *Firms maximise expected profits by hiring workers of the same expected skill.*

Proof See Appendix A.27.

Wages cannot be set according to individual skill, which is private information. As a consequence, each worker is paid exactly half of the firm's output. With this deliberately simple remuneration rule at hand, the expected wage rate of an individual worker with skill s then equals:

$$w(\bar{s}_{\ell t}, s) = A\bar{s}_{\ell t}s. \tag{5.20}$$

Interestingly, the wage of a worker with skill s not only depends on the worker's *own* skill s, but also on expected skill $\bar{s}_{\ell t}$ of the co-worker, which is assigned to the respective match. Although workers cannot observe the individual skill of their potential co-workers, the distribution of skills in both countries is known, such that expectations can be formed with regard to a potential co-worker's average skill $\bar{s}_{\ell t}$ with $\ell \in \{N, M, R\}$ and $t = 1, 2$.[29] Given this simple notion of workers' wages, the (return) migration decision of a forward-looking worker can be solved as a two-stage game (cf. Figure 5.9) with an initial emigration decision at age $t = 1$, and – in case of initial emigration – a later return decision at the age $t = 2$. We assume constant and equal periodical costs $c > 0$ for living abroad and solve the respective migration game by backward induction.[30]

As a natural starting point we begin with the return decision of worker s at the age $t = 2$, assuming that this worker emigrated abroad at the age $t = 1$. The forfeit wage gain $\Delta_r(s)$ for a worker s, who returns returns home at age $t = 2$ instead of staying a second and final period abroad, thereby amounts to:

$$\Delta_r(s) \equiv w(\bar{s}_{M2}, s) - c - w(\bar{s}_{R2}, s), \tag{5.21}$$

[29] Note that there are no returnees of age $t = 1$, hence, the subgroup $R \times 1$ does not exist.

[30] Periodical costs of living abroad in a new social and cultural environment without close relationships to friends and family back in the home country naturally arise for migrants who try to sustain familiar surroundings (cf. Sjaastad, 1962). Supportive evidence underpinning the importance of such costs comes for example from Gibson and McKenzie (2011), who use data on high-skilled return migration to three Pacific countries to show that return decisions are strongly linked to family and lifestyle reasons, rather than to the income differentials between source and destination countries.

and depends on the expected net wage $w(\bar{s}_{M2}, s) - c$ earned as a permanent migrant $(M \times 2)$ relative to the expected wage $w(\bar{s}_{R2}, s)$ earned as a returnee $(R \times 2)$. It is now straightforward to show that our model leads to self-selection of the most able initial emigrants $(M \times 1)$ into permanent migration. For this purpose we assume positive selection into initial emigration, which, as we will show below, indeed results at age $t = 1$. Thus, initial emigrants are assumed to have skills $s \geq \tilde{s}_m$ with $\tilde{s}_m \in (0, 1)$ denoting the skill of the least skilled emigrant at age $t = 1$. We may now consider some arbitrary cutoff ability $\tilde{s}_r \geq \tilde{s}_m$ that separates initial emigrants at age $t = 2$ into a group of high- and low-skilled individuals, which we precautionary label by $M \times 2$ and $R \times 2$, respectively. The average skills of the two groups immediately follow from the assumed uniform distribution and equal $\bar{s}_{M2} = (\tilde{s}_r + 1)/2 > \bar{s}_{R2} = (\tilde{s}_m + \tilde{s}_r)/2$. Substituting $\bar{s}_{M2}$ and $\bar{s}_{R2}$ into Eq. (5.21), we find that the expected wage gain from staying abroad equals $\Delta_r(s) = A(1 - \tilde{s}_m)s/2 - c$ and increases in individual skill s, such that incentives for staying abroad are high (low) for those workers with comparatively high (low) skills. Solving for the return cutoff, i.e. finding the marginal returnee $\tilde{s}_r$, to whom

$$\Delta_r(\tilde{s}_r) = \frac{A}{2}(1 - \tilde{s}_m)\tilde{s}_r - c \stackrel{!}{=} 0$$

applies, finally yields

$$\tilde{s}_r(\tilde{s}_m) = \frac{2\hat{c}}{1 - \tilde{s}_m}. \tag{5.22}$$

Intuitively, the mass of permanent migrants staying abroad is large if the (relative) costs $\hat{c} \equiv c/A$ of doing so are low. Given the inter-temporal structure of migration these costs – of course – must be weighted by the potential for permanent migration, i.e. the number of workers $1 - \tilde{s}_m$ who decided to emigrate in the first place at age $t = 1$.

Knowing how the return cutoff $\tilde{s}_r$ from Eq. (5.22) links to the initial emigration cutoff $\tilde{s}_m$, we now focus at workers' emigration decisions at age $t = 1$. Workers thereby take into account their later return decisions at age $t = 2$ and distinguish between three possible migration patterns, which we denote by (a), (b) and (c). The possible migration patterns are:

(a) $0 < \tilde{s}_m = \tilde{s}_r < 1 \quad \Rightarrow$ *permanent migration only,*

(b) $0 < \tilde{s}_m < \tilde{s}_r < 1 \quad \Rightarrow$ *temporary and permanent migration,*

(c) $0 < \tilde{s}_m < \tilde{s}_r = 1 \quad \Rightarrow$ *temporary migration only.*

In case (a) the emigration and the return cutoff coincide, i.e. $\tilde{s}_m = \tilde{s}_r$, such that all workers who stayed abroad at age $t = 1$ do the same at age $t = 2$. In case (b) only the best workers with skills $s \in [\tilde{s}_r, 1]$ stay for another period abroad, while workers with lower skills $s \in [\tilde{s}_m, \tilde{s}_r)$ return home at age $t = 2$. Finally, in case (c) with $\tilde{s}_r = 1$ everybody who emigrated at age $t = 1$ returns home at age $t = 2$. Taking into account these differences we can rank the average skills within the subgroups $\ell \times t$ with $\ell \in \{N, M, R\}$ and $t = 1, 2$ as follows:

$$\bar{s}_{\ell t} = \begin{cases} \bar{s}_{Nt} = \tilde{s}_m/2 < \bar{s}_{M1} = \bar{s}_{M2} = (\tilde{s}_m + 1)/2 & \text{if } (a), \\ \bar{s}_{Nt} = \tilde{s}_m/2 < \bar{s}_{R2} = (\tilde{s}_m + \tilde{s}_r)/2 < \bar{s}_{M1} = (\tilde{s}_m + 1)/2 < \bar{s}_{M2} = (\tilde{s}_r + 1)/2 & \text{if } (b), \\ \bar{s}_{Nt} = \tilde{s}_m/2 < \bar{s}_{M1} = \bar{s}_{R2} = (\tilde{s}_m + 1)/2 & \text{if } (c), \end{cases} \tag{5.23}$$

depending on which of the migration patterns (a), (b), or (c) results. The indifferent worker $\tilde{s}_m$ at age $t = 1$ then faces the following trade-off:

$$\Delta_m(\tilde{s}_m) = \begin{cases} w(\bar{s}_{M1}, \tilde{s}_m) + w(\bar{s}_{M2}, \tilde{s}_m) - 2c - w(\bar{s}_{N1}, \tilde{s}_m) - w(\bar{s}_{N2}, \tilde{s}_m) \stackrel{!}{=} 0 & \text{for } (a), \\ w(\bar{s}_{M1}, \tilde{s}_m) + w(\bar{s}_{R2}, \tilde{s}_m) - c - w(\bar{s}_{N1}, \tilde{s}_m) - w(\bar{s}_{N2}, \tilde{s}_m) \stackrel{!}{=} 0 & \text{for } (b), \\ w(\bar{s}_{M1}, \tilde{s}_m) + w(\bar{s}_{R2}, \tilde{s}_m) - c - w(\bar{s}_{N1}, \tilde{s}_m) - w(\bar{s}_{N2}, \tilde{s}_m) \stackrel{!}{=} 0 & \text{for } (c), \end{cases} \tag{5.24}$$

with $\Delta_m(s)$ denoting the expected lifetime income gain from going abroad, which for the indifferent initial emigrant $\tilde{s}_m$ by definition equals $\Delta_m(\tilde{s}_m) \stackrel{!}{=} 0$. Note that the opportunity cost of going abroad in all three cases materialise in form of the forfeit expected income stream $w(\bar{s}_{N1}, \tilde{s}_m) + w(\bar{s}_{N2}, \tilde{s}_m)$, that would result from domestic employment as a non-migrant ($N \times t$) at age $t = 1, 2$. On the contrary, when going abroad the indifferent emigrant $\tilde{s}_m$ at age $t = 1$ always earns an expected net wage of $w(\bar{s}_{M1}, \tilde{s}_m) - c$. At age $t = 2$ the indifferent emigrant's expected (net) wage then, however, depends on the underlying migration scenario. In case (a) with $0 < \tilde{s}_m = \tilde{s}_r < 1$ everybody including the indifferent emigrant $\tilde{s}_m$ stays abroad for a second (final) period and earns an expected net wage $w(\bar{s}_{M2}, \tilde{s}_m) - c$. In case ($b$) with $0 < \tilde{s}_m < \tilde{s}_r < 1$ only the best workers with $s \in [\tilde{s}_r, 1]$ stay permanently abroad, while the remaining workers $s \in [\tilde{s}_m, \tilde{s}_r)$, and in particular the indifferent emigrant $\tilde{s}_m$, return home to get employed at an expected wage $w(\bar{s}_{R2}, \tilde{s}_m)$. Finally, in case ($c$) everybody including the indifferent emigrant returns to home and earns an expected wage rate $w(\bar{s}_{R2}, \tilde{s}_m)$. Substituting $\bar{s}_{\ell t}$ with $\ell \in \{N, M, R\}$

and $t = 1, 2$ from Eq. (5.23) separately for the cases (a), (b) and (c) into Eq. (5.24) we obtain:

$$\Delta_m\left(\tilde{s}_m, \tilde{s}_r\right) = \begin{cases} A\tilde{s}_m - 2c \overset{!}{=} 0 & \text{for } (a), \\ A\left(1 + \tilde{s}_r\right)\tilde{s}_m/2 - c \overset{!}{=} 0 & \text{for } (b), \\ A\tilde{s}_m - c \overset{!}{=} 0 & \text{for } (c). \end{cases} \tag{5.25}$$

Note that in the extreme cases (a) and (c) the return margin $\tilde{s}_r$ is fixed, resulting either in no or in complete return migration, with $\tilde{s}_r = \tilde{s}_m$ or $\tilde{s}_r = 1$, respectively. In the intermediate case (b), on the contrary, the return margin $\tilde{s}_r \in (\tilde{s}_m, 1)$ is flexible and can be linked to the emigration cutoff $\tilde{s}_m$ through Eq. (5.22). Workers take this link into account when forming their emigration decisions at age $t = 1$ and anticipate that a lower (higher) return cutoff $\tilde{s}_r$ decreases (increases) the average skills $\bar{s}_{R2} = (\tilde{s}_m + \tilde{s}_r)/2 < \bar{s}_{M2} = (\tilde{s}_r + 1)/2$ within the groups of permanent migrants $(M \times 2)$ and returnees $(R \times 2)$ likewise.[31] Replacing $\tilde{s}_r$ in Eq. (5.25) by $\tilde{s}_r = 2\hat{c}/(1 - \tilde{s}_m)$ from Eq. (5.22) we can solve for the emigration cutoff $\tilde{s}_m^{\text{lf}}$ in the *laissez-faire* equilibrium:

$$\tilde{s}_m^{\text{lf}}\left(\hat{c}\right) = \begin{cases} 2\hat{c} & \text{for} \quad 0 \leq \hat{c} < \hat{c}_a^{\text{lf}} \quad \Leftrightarrow \quad (a) \;\; 0 < \tilde{s}_m^{\text{lf}} = \tilde{s}_r^{\text{lf}} < 1, \\ \dfrac{1 + 4\hat{c} - \sqrt{1 + 16\hat{c}^2}}{2} & \text{for} \quad \hat{c}_a^{\text{lf}} \leq \hat{c} < \hat{c}_b^{\text{lf}} \quad \Leftrightarrow \quad (b) \;\; 0 < \tilde{s}_m^{\text{lf}} < \tilde{s}_r^{\text{lf}} < 1, \\ \hat{c} & \text{for} \quad \hat{c}_b^{\text{lf}} \leq \hat{c} < \hat{c}_c^{\text{lf}} \quad \Leftrightarrow \quad (c) \;\; 0 < \tilde{s}_m^{\text{lf}} < \tilde{s}_r^{\text{lf}} = 1. \end{cases} \tag{5.26}$$

Finally, using $\tilde{s}_m^{\text{lf}}(\hat{c})$ from Eq. (5.26) to replace $\tilde{s}_m$ in Eq. (5.22) we can also solve for the return cutoff $\tilde{s}_r^{\text{lf}}(\hat{c})$ in the *laissez-faire* equilibrium:

$$\tilde{s}_r^{\text{lf}}\left(\hat{c}\right) = \begin{cases} 2\hat{c} & \text{for} \quad 0 \leq \hat{c} < \hat{c}_a^{\text{lf}} \quad \Leftrightarrow \quad (a) \;\; 0 < \tilde{s}_m^{\text{lf}} = \tilde{s}_r^{\text{lf}} < 1, \\ \dfrac{4\hat{c}}{1 - 4\hat{c} + \sqrt{1 + 16\hat{c}^2}} & \text{for} \quad \hat{c}_a^{\text{lf}} \leq \hat{c} < \hat{c}_b^{\text{lf}} \quad \Leftrightarrow \quad (b) \;\; 0 < \tilde{s}_m^{\text{lf}} < \tilde{s}_r^{\text{lf}} < 1, \\ 1 & \text{for} \quad \hat{c}_b^{\text{lf}} \leq \hat{c} < \hat{c}_c^{\text{lf}} \quad \Leftrightarrow \quad (c) \;\; 0 < \tilde{s}_m^{\text{lf}} < \tilde{s}_r^{\text{lf}} = 1. \end{cases} \tag{5.27}$$

Together $\tilde{s}_m^{\text{lf}}(\hat{c})$ from Eq. (5.26) and $\tilde{s}_r^{\text{lf}}(\hat{c})$ from Eq. (5.27) provide a comprehensive description of the intertemporal migration pattern. In particular we find that for rising (relative) costs $\hat{c} \in [0, \infty)$ of staying abroad the migration patterns (a), (b) and (c) become relevant in increasing

[31] The strategic link between initial emigration and later return decisions is explored in more detail in Section 5.8.2, where we introduce a discounting factor $\delta \in [0, 1]$, which allows for an asymmetric weighting of workers' payoffs in the migration decisions at age $t = 1$ and $t = 2$, respectively.

order, with the parameter spaces corresponding to each of these cases being separated by the cost thresholds $\hat{c}_a^{\text{lf}} \leq \hat{c}_b^{\text{lf}} \leq \hat{c}_c^{\text{lf}}$. The cost thresholds $\hat{c}_a^{\text{lf}}$ and $\hat{c}_b^{\text{lf}}$ thereby follow immediately from the definitions of the limiting cases (a) and (c), restricting case (b) from above and below. Focusing on case (b) and using the corresponding expressions $\tilde{s}_m^{\text{lf}}(\hat{c})$ and $\tilde{s}_r^{\text{lf}}(\hat{c})$ from Eqs. (5.26) and (5.27), we find that $\tilde{s}_m^{\text{lf}}(\hat{c}_a^{\text{lf}}) \stackrel{!}{=} \tilde{s}_r^{\text{lf}}(\hat{c}_a^{\text{lf}})$ and $\tilde{s}_m^{\text{lf}}(\hat{c}_b^{\text{lf}}) \stackrel{!}{=} 1$ imply $\hat{c}_a^{\text{lf}} = 0$ and $\hat{c}_b^{\text{lf}} = 1/3$, respectively. Finally, to identify the critical cost level $\hat{c}_c^{\text{lf}}$ which separates case (c) from an equilibrium without any migration we focus on $\tilde{s}_m^{\text{lf}}(\hat{c}_c^{\text{lf}}) \stackrel{!}{=} 1$ from Eq. (5.26) for case (c), and find $\hat{c}_c^{\text{lf}} = 1$. Proposition 5.6.2 summarises the results:

Proposition 5.6.2 *At prohibitive costs $\hat{c} \geq \hat{c}_c^{lf}$ no migration occurs. In case (c) for high but not prohibitively high costs $\hat{c} \in [\hat{c}_b^{lf}, \hat{c}_b^{lf})$ high-skilled workers $s \in [\tilde{s}_m^{lf}, 1]$ migrate temporary abroad, while low-skilled workers $s \in [0, \tilde{s}_m^{lf})$ stay at home. Finally, in the low-cost scenario (b) with $\hat{c} \in (0, \hat{c}_b^{lf})$ high-skilled workers $s \in [\tilde{s}_r^{lf}, 1]$ emigrate permanently abroad, medium-skilled workers $s \in [\tilde{s}_m^{lf}, \tilde{s}_r^{lf})$ migrate temporary, and low-skilled workers $s \in [0, \tilde{s}_m^{lf})$ do not migrate at all.*[32]

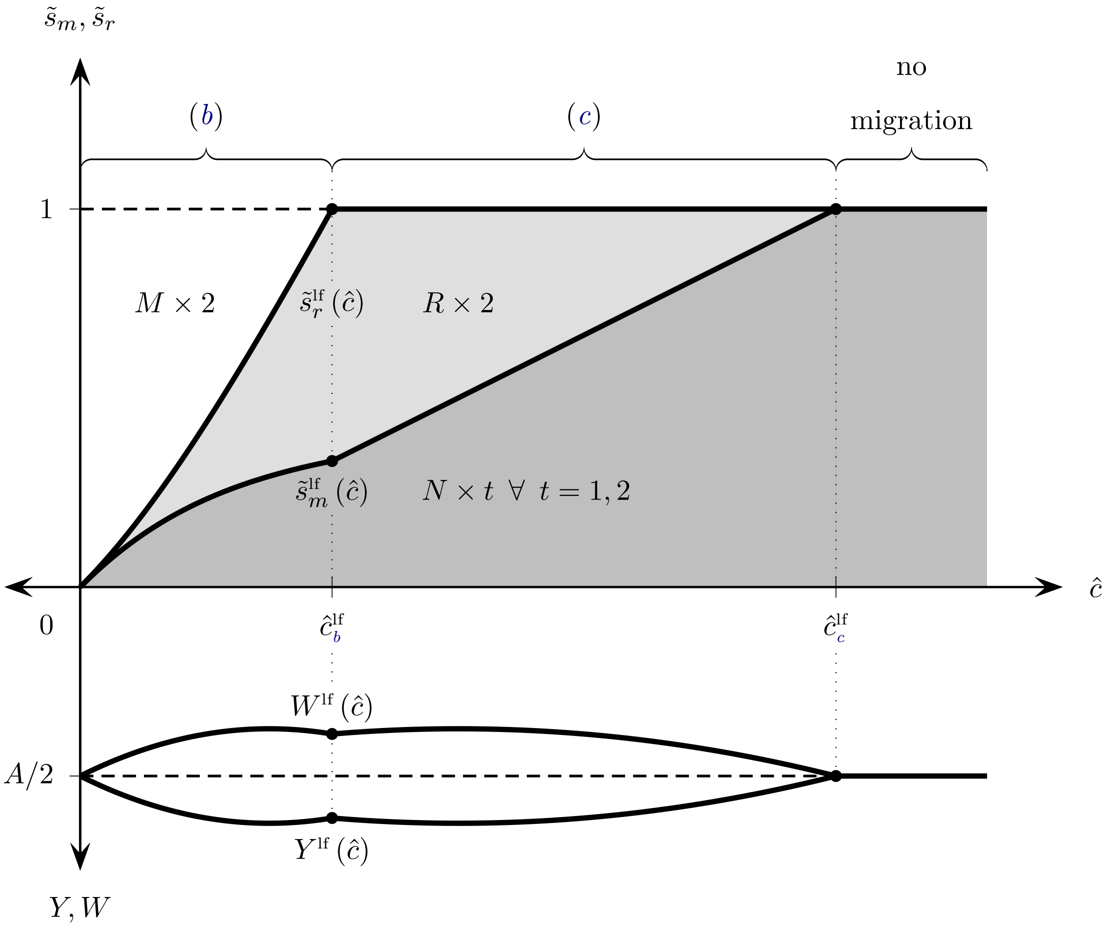
Proof Analysis and formal discussion in the text.

To understand the migration pattern in Proposition 5.6.2 it is important to realise that high-skilled workers are the only ones that can afford the costs c of staying abroad. A costly stay abroad, hence, acts as a signalling device for these workers and gives them the opportunity to indicate their high but otherwise unobservable skills towards potential employers. Firms, when making their hiring decisions, take individual migration histories as an easy-to-verify signal into account and form, in line with Lemma 5.6.1, more efficient and better paid matches at the labour market. The wage premium resulting from the increased quality of matches is then what gives workers an incentive to signal their skills in the first place.

[32] Proposition 5.6.2 in combination with Lemma 5.6.1 implies that natives and migrants are perfectly segregated at the workplace. While this extreme implication is counterfactual of course, Hellerstein and Neumark (2008); Andersson, García-Pérez, Haltiwanger, McCue, and Sanders (2010); Aslund and Skans (2010); Glitz (2012) find that there is indeed a substantial degree of workplace segregation between natives and migrants in the US, Sweden, and Germany. As outlined in Section 5.4.1 our migration mechanism can plausibly replicate this outcome, once we allow individual skills to be imperfectly observable (instead of unobservable).

Since, workers are allowed to stay abroad for more than just one period, we find that signalling decisions are linked through time and give rise to the inter-temporal pattern of initial emigration and (eventually) later return migration. We depict this pattern in the upper quadrant of Figure 5.10. Figure 5.10 thereby distinguishes between the groups of non-migrants $N \times t$

Figure 5.10: *The laissez-faire equilibrium*



at age $t = 1, 2$ (dark grey area), return migrants $R \times 2$ (light grey area), and permanent migrants $M \times 2$ (white area), whose relative size crucially depends on the underlying migration pattern and, thus, on the (relative) costs $\hat{c}$ of staying abroad. If these costs are prohibitively high, i.e. $\hat{c} \geq \hat{c}_c^{\text{lf}}$, not even the most high-skilled workers find it optimal to signal their skills by staying at least one period abroad. On the contrary, for high but not prohibitively high costs $\hat{c} \in [\hat{c}_b^{\text{lf}}, \hat{c}_c^{\text{lf}})$ migration pattern (c) results. In this case the most high-skilled workers with

$s \in [\tilde{s}_m^{\mathrm{lf}}, 1]$ emigrate abroad at age $t = 1$, thereby separating themselves from their low-skilled counterparts with $s \in [0, \tilde{s}_m^{\mathrm{lf}})$. Then, at age $t = 2$, all those who initially emigrated at age $t = 1$ return back home. The complete reversal of initial migration incentives is caused by the *now* (at age $t = 2$) prohibitive costs $\hat{c}$ of staying another period abroad, which render the use of a repeated stay abroad as signalling device suboptimal, given the already achieved separation from the group of the most low-skilled workers (comprising the non-migrants $N \times t \;\; \forall \;\; t = 1, 2$). Let us now consider migration pattern (*b*), which results for low (relative) costs $\hat{c} \in (0, \hat{c}_b^{\mathrm{lf}})$. As in the previous case all workers with high skills $s \in [\tilde{s}_m^{\mathrm{lf}}, 1]$ use initial emigration at age $t = 1$ as a signal to achieve a separation from their low-skilled counterparts with low skills $s \in [0, \tilde{s}_m^{\mathrm{lf}})$. However, then (at age $t = 2$) only the medium-skilled workers with skill $s \in [\tilde{s}_m^{\mathrm{lf}}, \tilde{s}_r^{\mathrm{lf}})$ return home, while the most high-skilled workers with skills $s \in [\tilde{s}_r^{\mathrm{lf}}, 1]$ stay abroad for a second and final period. High-skilled workers thereby generate an effective, since costly, signal which can be used to tell apart high-skilled permanent migrants ($M \times 2$) from medium skilled returnees ($R \times 2$). Permanent migrants ($M \times 2$) are rewarded for their signalling efforts by more efficient matches and, hence, higher average wages. On the contrary, returnees ($R \times 2$) experience a decline in the quality of their potential co-workers and, thus, a fall in average wages. However, at the same time they also save on the costs c, which overcompensates their expected wage losses and, hence, tips the scales in favour of return migration. Finally, migration pattern (*a*), featuring permanent migration only, never results. To understand this outcome, it is helpful to compare the marginal emigrant's emigration and return decisions. By emigrating abroad under costs c when still being young (i.e. at age $t = 1$) the marginal emigrant $\tilde{s}_m^{\mathrm{lf}}$ gets access to "good" matches within the group of initial emigrants ($M \times 1$) and avoids "bad" matches within the group of non-migrants ($N \times 1$). At age $t = 2$, staying abroad at the same costs c ensures the continued access to "good" matches within the group of (now) permanent emigrants ($M \times 2$). However, even when not bearing the costs c there is no danger of ending up in a "bad" match with low-skilled non-migrants ($N \times 2$) as long as being identifiable as a returnee ($R \times 2$). Thus, for the critical emigrant $\tilde{s}_m^{\mathrm{lf}}$ the gain from using a costly stay abroad as signalling device is (by definition) zero at age $t = 1$ and even smaller at age $t = 2$. Hence, for any costs $\hat{c} > 0$, which are non-decreasing in workers' age, migration pattern (a) with $0 < \tilde{s}_m^{\text{lf}} = \tilde{s}_r^{\text{lf}} < 1$ cannot exist.[33]

Taking stock, our simple model is able to generate a rich pattern of temporary *and* permanent two-way migration between two ex post and ex ante identical countries. Our framework thereby offers an explanation for the balance in temporary *and* permanent bilateral migration stocks documented in Section 5.5. As a key element of our model individual emigration *and* return decisions are jointly derived from the (repeated) use of costly stays abroad as signalling devices (cf. Spence, 1973) and, hence, do not follow from presupposed country asymmetries as usually assumed in the literature on temporary migration (cf. Djajic and Milbourne, 1988; Djajic, 1989; Dustmann, 2001; Dustmann and Weiss, 2007; Dustmann, Fadlon, and Weiss, 2011). Self-selection of workers into (return) migration within our framework thereby is based on skills that are private information, which we do not want to confuse with selection on observables as analysed by Borjas (1987, 1991); Borjas and Bratsberg (1996). Supportive evidence for selection patterns that are based on characteristics, which usually are regarded as *unobservable* comes from Dostie and Léger (2009), who focus on the inter-provincial migration of Canadian physicians and are able to decompose physicians' earnings into an observable and an unobservable component. Thereby it turns out that within this narrowly defined group of high-skilled workers positive selection into migration is driven by unobservable rather than by observable characteristics.[34]

5.7 Welfare effects of temporary migration

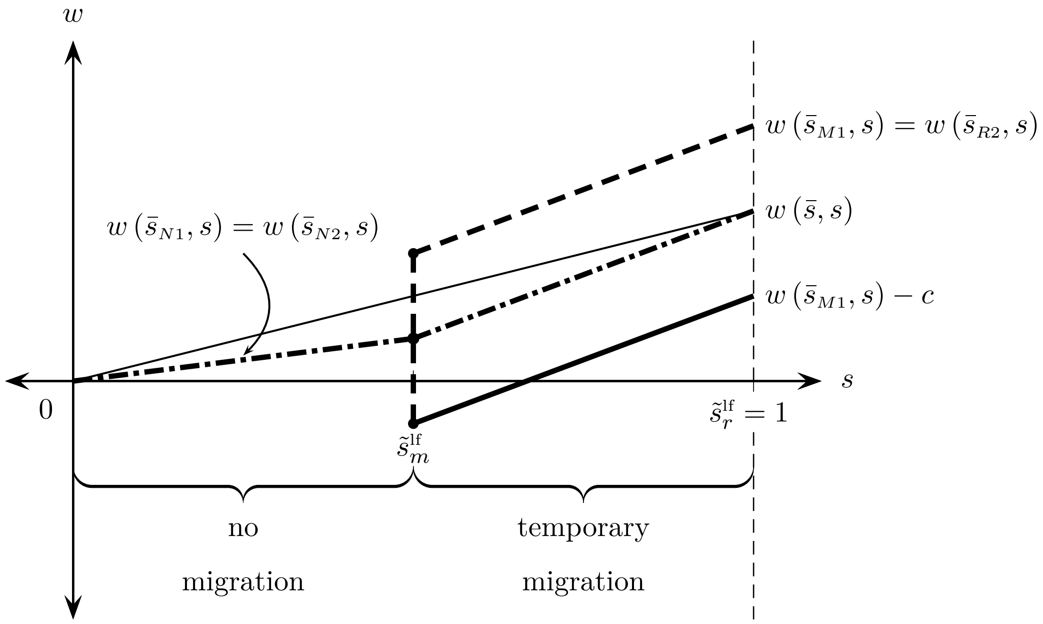
With the simple migration pattern in Proposition 5.6.2 at hand, we are now equipped for an exploration of the individual welfare effects that initial emigration in combination with possible later return migration may have. As a natural welfare measure at the individual level we focus on workers' expected (net) lifetime income, i.e. expected wages at age $t = 1, 2$ less the migration

[33]In Section 5.8.3 we allow the costs c of staying abroad to decline with the workers' age t and show for which precise parameter range the migration pattern (a) exists.

[34]Using data on (return) migration between Finland and Sweden Rooth and Saarela (2007) find no selection with respect to unobservable skills, which they attribute to the fact that both countries are too similar in various aspects.

costs c, if applicable. We plot the economy's wage profiles separately for the cases (c) and (b) in the Figures 5.11 and 5.12 and begin our analysis with the high-cost case (c). Figure 5.11 depicts the expected gross wages, $w(\bar{s}_{Nt}, s)\ \ \forall\ t = 1, 2$ for non-migrants (thick dot-dashed line) with $s \in [0, \tilde{s}_m^{\mathrm{lf}})$ as well as the expected gross wages $w(\bar{s}_{M1}, s) = w(\bar{s}_{R2}, s)$ for initial emigrants and later returnees (thick dashed line) with $s \in [\tilde{s}_m^{\mathrm{lf}}, 1]$. For comparison the expected autarky wage profile $w(\bar{s}, s)$ is depicted as a thin solid line. Since temporary migrants signal their skills

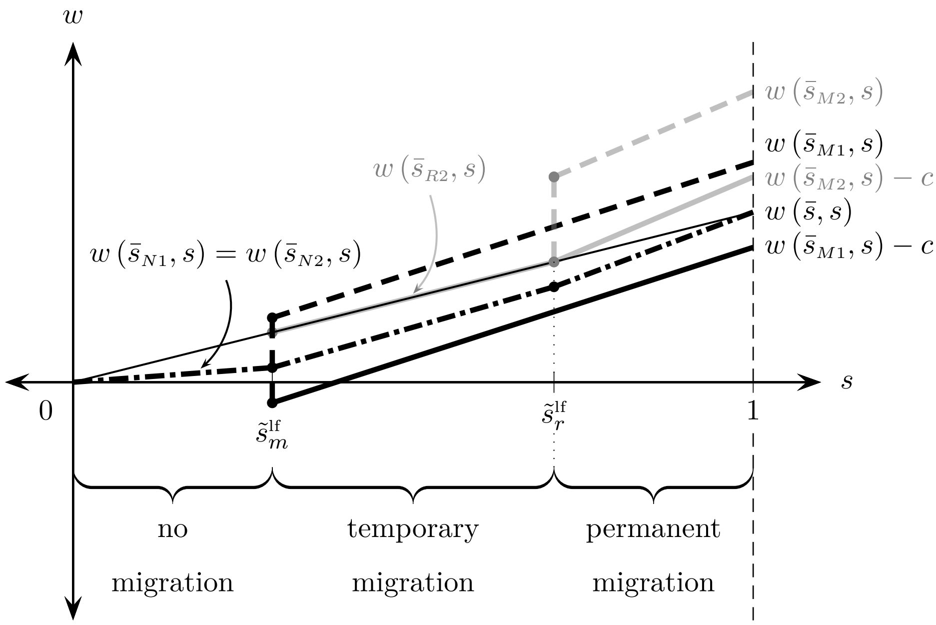
Figure 5.11: *Wage distribution for migration pattern (c)*



by staying abroad, we find expected gross wage gains for initial emigrants and later returnees. Non-migrants, on the contrary, suffer expected wage losses from the decreasing quality of co-workers within their group. While the wage losses for non-migrants are directly welfare relevant, wage gains for temporary migrants must be set off against the costs c of staying abroad at age $t = 1$. Taking this signalling costs into account we find that expected net wages for initial emigrants $w(\bar{s}_{M1}, s) - c$ (thick solid line) are considerably smaller than the respective autarky wages $w(\bar{s}, s)$. Of course, these income losses at age $t = 1$ must be seen in relation to the return premium $w(\bar{s}_{R2}, s) - w(\bar{s}, s) > 0$ that temporary migrants earn when being back in their home economy. Averaging net wages, $w(\bar{s}_{M1}, s) - c$ and $w(\bar{s}_{R2}, s)$, over workers' age (thick dot-dashed line) we find that both non-migrants and temporary migrants do *not* gain relative to an equilibrium without migration.

The economy's wage profile for the low-cost scenario (b) is depicted in Figure 5.12. Unlike before (in Figure 5.11) we now have to distinguish between three subgroups of workers: Non-migrants with $s \in [0, \tilde{s}_m^{\text{lf}})$, temporary migrants with $s \in [\tilde{s}_m^{\text{lf}}, \tilde{s}_r^{\text{lf}})$, and permanent migrants for which $s \in [\tilde{s}_r^{\text{lf}}, 1]$. Non-migrants again suffer in terms of low expected wages $w(\bar{s}_{Nt}, s)\ \forall\ t = 1, 2$

Figure 5.12: *Wage distribution for migration pattern (b)*



(thick black dot-dashed line) as the quality of co-workers within their group falls if the most high-skilled workers emigrate abroad. On the contrary, initial emigrants benefit from the signalling effect of staying abroad and earn higher expected gross wages $w(\bar{s}_{M1}, s)$ (thick black dashed line) than in an equilibrium without migration (thin solid line). However, once the signalling costs c are taken into account emigrants' expected net wages $w(\bar{s}_{M1}, s) - c$ (thick black solid line) are below the respective autarky wage profile $w(\bar{s}, s)$ (thin solid line). At age $t = 2$ temporary migrants earn expected wages $w(\bar{s}_{R2}, s)$ (thick grey solid line) above $w(\bar{s}_{M1}, s) - c$ and below $w(\bar{s}_{M1}, s)$.[35] The reason for this outcome is simple: On the one hand, temporary migrants save

[35] Note that returnees' expected wages $w(\bar{s}_{R2}, s)$ can be below or above the respective autarky wages $w(\bar{s}, s)$.

on costs c by returning home, while, on the other hand, the most high-skilled initial emigrants decide to stay (permanently) abroad, thereby reducing the quality of available co-workers within the group of returnees. Finally, permanent migrants signal their high skills through a repeated and final stay abroad. As a consequence they enjoy higher expected gross wages $w(\bar{s}_{M2}, s)$ (thick grey dashed line), which, at least for the most high-skilled among the permanent migrants, are always above the respective autarky wage, even when taking the signalling costs c into account (thick grey solid line). To compare individual welfare in an equilibrium with and without migration we again take the average of (net) wages over time (thick black dot-dashed line) and find that in a migration equilibrium *no* worker is better off than under autarky.

Interestingly, aggregate (net) income not only is smaller than under autarky, it also is more unequal distributed. To see this note that the expected income profile over age $t = 1, 2$ in a migration equilibrium (thick black dot-dashed line) is kinked once in case (c) and twice in case (b). As a consequence the (expected) returns to skill increase slower (faster) at the lower (upper) end of the skill distribution, which renders the corresponding income distribution more unequal than in an equilibrium without migration.

In order to compute aggregate welfare we start out from aggregate production

$$Y = \sum_{t} \sum_{\ell} \int_{s \in \ell \times t} w\left(\bar{s}_{\ell t}, s\right) \, ds, \tag{5.28}$$

which in a zero-profit equilibrium is defined as the sum of workers' expected wages over all skill levels s within the subgroups $\ell \times t$ with $\ell \in \{N, M, R\}$ and $t = 1, 2$. Using the definition of $w(\bar{s}_{\ell t}, s)$ in Eq. (5.20) in combination with $\bar{s}_{\ell t}$ from Eq. (5.23) and replacing $\tilde{s}_m$ and $\tilde{s}_r$ by $\tilde{s}_m^{\mathrm{lf}}$ and $\tilde{s}_r^{\mathrm{lf}}$ from Eqs. (5.26) and (5.27) we can solve separately for the cases (b) and (c) as well as

For expositional purpose Figure 5.12 depicts the knife-edge case $w(\bar{s}_{R2}, s) = w(\bar{s}, s)$. As a sufficient condition for a return premium as documented in Co, Gang, and Yun (2000), Barrett and O'Connell (2001), Barrett and Goggin (2010) the costs of staying abroad should not be too low, i.e. $\hat{c} > 1/4$, as otherwise too many high-skilled workers decide in favour of permanent migration.

for the autarky case:

$$Y^{\mathrm{lf}}(\hat{c}) = \begin{cases} A/2 + A(1-2\hat{c})\,\hat{c} & \text{for} \quad \hat{c} \in [\hat{c}_a^{\mathrm{lf}}, \hat{c}_b^{\mathrm{lf}}) \quad \Leftrightarrow \quad (b)\;\; 0 < \tilde{s}_m^{\mathrm{lf}} < \tilde{s}_r^{\mathrm{lf}} < 1, \\ A/2 + A(1-\hat{c})\,\hat{c} & \text{for} \quad \hat{c} \in [\hat{c}_b^{\mathrm{lf}}, \hat{c}_c^{\mathrm{lf}}) \quad \Leftrightarrow \quad (c)\;\; 0 < \tilde{s}_m^{\mathrm{lf}} < \tilde{s}_r^{\mathrm{lf}} = 1, \\ A/2 & \text{for} \quad \hat{c} \geq \hat{c}_c^{\mathrm{lf}} \quad \Leftrightarrow \quad \tilde{s}_m^{\mathrm{lf}} = \tilde{s}_r^{\mathrm{lf}} = 1. \end{cases} \tag{5.29}$$

Plotting $Y^{\mathrm{lf}}(\hat{c})$ in the lower quadrant of Figure 5.10 we find that aggregate production is higher in any migration equilibrium than in an equilibrium without migration. Production gains from migration thereby arise from the more efficient matches among workers, which firms can form according to Lemma 5.6.1 when taking into account the information embodied in workers' (return) migration decisions.

Aggregate welfare can now be derived from Eq. (5.28) by taking into account the costs c associated with staying abroad at age $t = 1$ and $t = 2$, respectively

$$W = Y - \sum_t \int_{s \in M \times t} c \; ds. \tag{5.30}$$

Substituting $Y^{\mathrm{lf}}(\hat{c})$ from Eq. (5.29), while aggregating the costs c over all initial migrants $M \times 1$ with $s \in [\tilde{s}_m^{\mathrm{lf}}, 1]$ as well as over all permanent migrants $(M \times 2)$ with $s \in [\tilde{s}_r^{\mathrm{lf}}, 1]$ we obtain:

$$W^{\mathrm{lf}}(\hat{c}) = \begin{cases} A/2 - A(1-2\hat{c})\,\hat{c} & \text{for} \quad \hat{c} \in [\hat{c}_a^{\mathrm{lf}}, \hat{c}_b^{\mathrm{lf}}) \quad \Leftrightarrow \quad (b)\;\; 0 < \tilde{s}_m^{\mathrm{lf}} < \tilde{s}_r^{\mathrm{lf}} < 1, \\ A/2 - A(1-\hat{c})\,\hat{c} & \text{for} \quad \hat{c} \in [\hat{c}_b^{\mathrm{lf}}, \hat{c}_c^{\mathrm{lf}}) \quad \Leftrightarrow \quad (c)\;\; 0 < \tilde{s}_m^{\mathrm{lf}} < \tilde{s}_r^{\mathrm{lf}} = 1, \\ A/2 & \text{for} \quad \hat{c} \geq \hat{c}_c^{\mathrm{lf}} \quad \Leftrightarrow \quad \tilde{s}_m^{\mathrm{lf}} = \tilde{s}_r^{\mathrm{lf}} = 1. \end{cases} \tag{5.31}$$

Plotting $W^{\mathrm{lf}}(\hat{c})$ again in the lower quadrant of Figure 5.10 we find that in any migration equilibrium aggregate welfare is lower than in an equilibrium without migration. Proposition 5.7.1 summarises the results.

Proposition 5.7.1 *In a migration equilibrium aggregate welfare (production) is lower (higher) than in an equilibrium without migration. Moreover, in a migration equilibrium no worker is better off than under autarky.*

Proof Analysis in the text and formal discussion in Appendix A.28.

The reason behind the welfare loss described in Proposition 5.7.1 is a *negative* signalling externality, which renders initial and permanent emigration *too* attractive in the *laissez-faire* equilibrium. As a consequence the production gains from the more efficient matching of workers at the labour market are eaten up by the wasteful costs c of staying abroad, which are associated with excessive initial and permanent emigration at age $t = 1$ and $t = 2$. Thereby, the negative signalling externality can best be seen by means of a thought experiment, in which initial (permanent) emigration occurs sequentially, in the order of decreasing ability of emigrants: Every initial (permanent) emigrant, apart from the most skilled one, by going (staying) abroad lowers the average skill within the group of non-migrants (returnees) and within the group of initial (permanent) emigrants. For the groups of non-migrants (returnees) this is the case because the critical initial (permanent) emigrant leaves this group as the most high-skilled worker. For the groups of initial (permanent) migrants this is the case because the critical initial (permanent) emigrant enters this group as the most low-skilled worker. Thus, whenever a worker decides to stay abroad he inflicts losses on the wages of everyone else. These social costs are rationally ignored in individual emigration decisions and, hence, lead to excessive initial and permanent emigration.

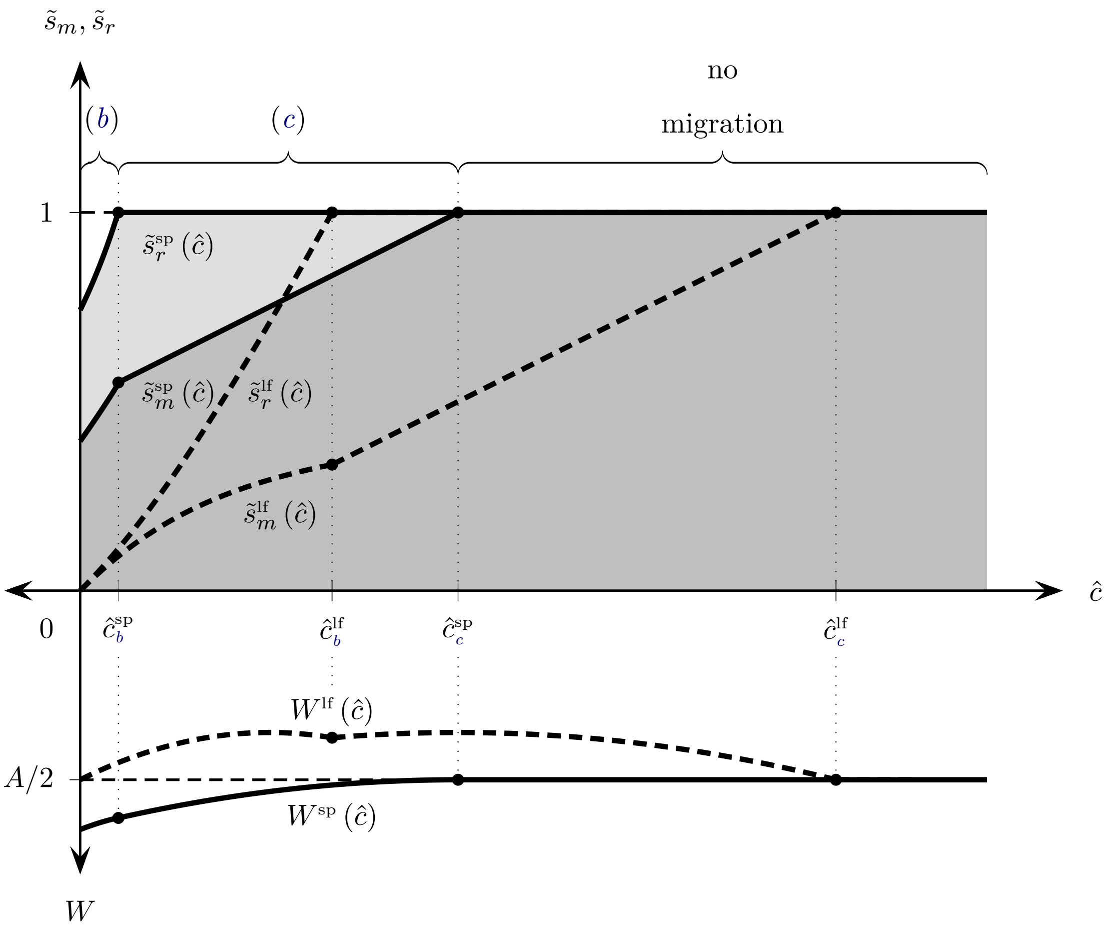
In order to characterise the optimal migration policy in the above model we assume, similar to Benhabib and Jovanovic (2012), an omniscient, global social planner, which can freely choose the initial emigration cutoff $\tilde{s}_m$ in combination with the return cutoff $\tilde{s}_r$. The social planner thereby ignores individual (return) migration incentives which link $\tilde{s}_m^{\text{lf}}$ and $\tilde{s}_r^{\text{lf}}$ to $\hat{c}$ in the *laissez-faire* equilibrium and maximises instead aggregate welfare in Eq. (5.30) with respect to $\tilde{s}_m$ and $\tilde{s}_r$. By deciding on a specific combination of $\tilde{s}_m^{\text{sp}}(\hat{c}), \tilde{s}_r^{\text{sp}}(\hat{c}) \in [0, 1]$ the social planner not only determines whether the two economies end up in migration scenario (*a*), (*b*), or (*c*) instead of autarky, but also links the extent of initial and permanent emigration (captured by $\tilde{s}_m^{\text{sp}}(\hat{c})$ and $\tilde{s}_r^{\text{sp}}(\hat{c})$) in an socially optimal way to the underlying (relative) cost $\hat{c}$. We depict the socially optimal emigration and return cutoffs, $\tilde{s}_m^{\text{sp}}(\hat{c})$ and $\tilde{s}_r^{\text{sp}}(\hat{c})$, in Figure 5.13 along with the implied (optimal) level of aggregate welfare $W^{\text{sp}}(\hat{c})$ and discuss the results in Proposition 5.7.2.

Proposition 5.7.2 *The socially optimal level of initial and permanent emigration is strictly lower than in the laissez-faire equilibrium, if the latter features positive migration levels. For*

$\hat{c} < \hat{c}_c^{sp} = 1/2$ ($\hat{c} < \hat{c}_b^{sp} \approx 1/20$) *the socially optimal level of initial (permanent) emigration is strictly positive.*

Proof Analysis in the text and formal discussion in Appendix A.29.

Figure 5.13: *The social-planner equilibrium*



The social planner internalises the negative external effect of initial and permanent emigration and chooses higher cutoffs $\tilde{s}_k^{sp}(\hat{c}) > \tilde{s}_k^{lf}(\hat{c})\ \forall\ k = m, r$ than in the *laissez-faire* equilibrium. Thereby the social planner not necessarily enforces a non-migration equilibrium. In particular at low costs $\hat{c} < \hat{c}_b^{sp}$ ($\hat{c} < \hat{c}_c^{sp}$) we find that the aggregate production gains from improved matching at the labour market exceed the associated costs of staying (repeatetly) abroad. As a consequence migration pattern (b) featuring temporary *and* permanent migration is implemented for low costs $0 < \hat{c} < \hat{c}_b^{\text{sp}}$, while the temporary-migration-only pattern (c) with $\tilde{s}_m^{\text{sp}}(\hat{c}) = \frac{1}{2} + \tilde{s}_m^{\text{lf}}(\hat{c})$ follows for high, but not prohibitively high costs $\hat{c}_b^{\text{sp}} \leq \hat{c} < \hat{c}_c^{\text{sp}}$. Migration pattern ($a$), on the contrary, is never chosen by the social planner and it is easy to see why: Each matching of skills within the group of high-skilled permanent migrants ($M \times t$) and the group of low-skilled non-migrants ($N \times t$) in the permanent-migration-only equilibrium (a) can always be replicated by an identical matching result within the groups of high-skilled temporary migrants ($M \times 1$ or $R \times 2$) and low-skilled non-migrants ($N \times t$) in the temporary-migration-only equilibrium (c). The only difference between both cases are the costs c, which accrue once for all temporary migrants in case (c), but twice for all permanent migrants in case (a).

We now show that the social-planner equilibrium can alternatively be implemented by a carefully chosen combination of emigration tax $\tau_m > 0$ and return subsidy $\tau_r < 0$. Thereby we assume that any surplus in tax revenues is redistributed equally among all nationals, irrespective of their residence, and, thus, has no effect individual (return) migration decisions. What *both* countries then care about is emigration not immigration. Intuitively, given that migrants and natives do not interact in the labour market (cf. Lemma 5.6.1), immigration has no effect on natives' wages and, hence, no impact on nationals' well-being. At the same time, however, temporary (permanent) emigration is associated with an negative external effect, depressing nationals' wages in the *laissez-faire* equilibrium at home and abroad. Hence, to replicate the social-planner equilibrium both countries will set a combination of emigration tax τ_m^{sp} and return subsidy τ_r^{sp}, which leads to $\tilde{s}_k^{\text{lf}}(\hat{c}_1, \hat{c}_2) = \tilde{s}_k^{\text{sp}}(\hat{c})$, with $\hat{c}_1 = \hat{c} + \hat{\tau}_m$ and $\hat{c}_2 = \hat{c} + \hat{\tau}_r$, whereas $\hat{\tau}_k \equiv \tau_k / A \ \forall \ k = m, r$. Thereby it is important to note that for $\tau_m \neq \tau_r$ the cost of living abroad c_t at age $t = 1, 2$ no longer are the same. Allowing these cost to be age-dependent and going through the same steps as in Section 5.6 we can generalise Eqs. (5.26) and (5.27) to:

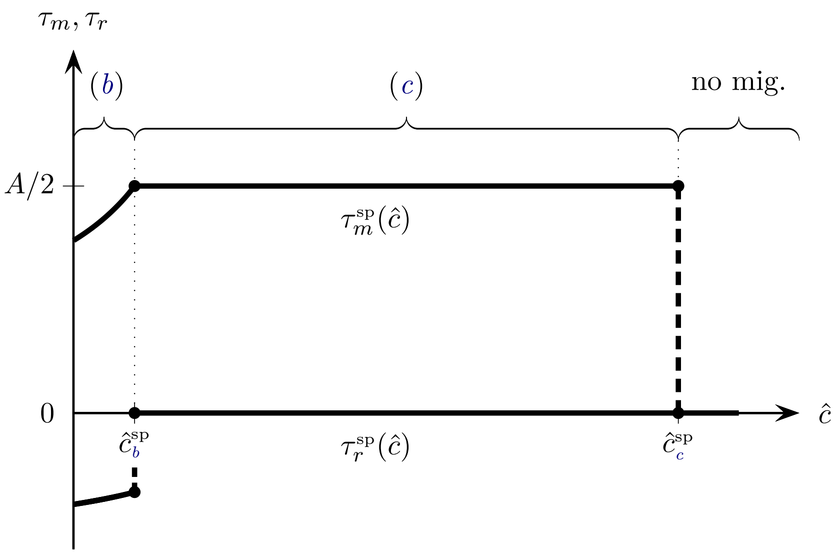
$$\tilde{s}_m^{\text{lf}}(\hat{c}_1, \hat{c}_2) = \begin{cases} 2\hat{c}_1 & \text{for } (a)\ 0 < \tilde{s}_m^{\text{lf}} = \tilde{s}_r^{\text{lf}} < 1, \\ \dfrac{1 + 2\hat{c}_1 + 2\hat{c}_2 - \sqrt{(1 + 2\hat{c}_1 + 2\hat{c}_2)^2 - 8\hat{c}_2}}{2} & \text{for } (b)\ 0 < \tilde{s}_m^{\text{lf}} < \tilde{s}_r^{\text{lf}} < 1, \\ \hat{c}_1 & \text{for } (c)\ 0 < \tilde{s}_m^{\text{lf}} < \tilde{s}_r^{\text{lf}} = 1, \end{cases} \tag{5.26'}$$

and

$$\tilde{s}_r^{\text{lf}}(\hat{c}_1,\hat{c}_2) = \begin{cases} 2\hat{c}_1 & \text{for } (a)\ 0 < \tilde{s}_m^{\text{lf}} = \tilde{s}_r^{\text{lf}} < 1, \\ \dfrac{4\hat{c}_2}{1-2\hat{c}_1-2\hat{c}_2+\sqrt{(1+2\hat{c}_1+2\hat{c}_2)^2-8\hat{c}_2}} & \text{for } (b)\ 0 < \tilde{s}_m^{\text{lf}} < \tilde{s}_r^{\text{lf}} < 1, \\ 1 & \text{for } (c)\ 0 < \tilde{s}_m^{\text{lf}} < \tilde{s}_r^{\text{lf}} = 1. \end{cases} \tag{5.27$'$}$$

Using $\tilde{s}_k^{\text{lf}}(\hat{c}_1,\hat{c}_2) = \tilde{s}_k^{\text{sp}}(\hat{c})$, with $k = m, r$ and $\tilde{s}_k^{\text{lf}}(\hat{c}_1,\hat{c}_2)$ from Eqs. (5.26′) and (5.27′) we can now solve for the optimal emigration tax $\tau_m^{\text{sp}}(\hat{c})$ and the optimal return subsidy $\tau_r^{\text{sp}}(\hat{c})$. As we would expect optimal migration policies differ between the migration scenarios (b) and (c), which we jointly depict in Figure 5.14. For the high-cost case (c) it follows immediately from $\tilde{s}_m^{\text{sp}}(\hat{c}) = \frac{1}{2} +$

Figure 5.14: *Optimal migration policies*



$\tilde{s}_m^{\text{lf}}(\hat{c})$ that both countries raise a constant emigration tax of $\tau_m(\hat{c}) = A/2\ \forall\ \hat{c} \in [\hat{c}_b^{\text{sp}}, \hat{c}_c^{\text{sp}})$ in order to prevent excessive temporary migration at age $t = 1$. In the low-cost case (b) for $\hat{c} \in [\hat{c}_a^{\text{sp}}, \hat{c}_b^{\text{sp}})$ migration policies are more complex: To achieve the socially optimal low levels of temporary *and* permanent migration both sending countries tax emigration at age $t = 1$ and subsidise return migration at age $t = 2$. The socially optimal combination of emigration tax $\tau_m^{\text{sp}}(\hat{c})$ and return subsidy $\tau_r^{\text{sp}}(\hat{c})$ depicted in Figure 5.14 thereby not only replicates the social-planner equilibrium in Figure 5.13, but also generates a surplus in tax revenues $\tau_m^{\mathrm{sp}}(\hat{c})[1-\tilde{s}_m^{\mathrm{sp}}(\hat{c})]+\tau_r^{\mathrm{sp}}(\hat{c})[1-\tilde{s}_r^{\mathrm{sp}}(\hat{c})] > 0$, which, if distributed equally among all nationals, mitigates the increased income inequality in the migration equilibrium. The same logic of course also applies to the high-cost case (c) in which no expenditures on return subsidies have to be covered. Proposition 5.7.3 summarises the results:

Proposition 5.7.3 *As optimal migration policies both countries choose symmetric emigration taxes $\tau_m^{sp}(\hat{c})$ if the costs $\hat{c} \in [\hat{c}_b^{sp}, \hat{c}_c^{sp})$ are high, and a symmetric combination of emigration taxes $\tau_m^{sp}(\hat{c})$ and return subsidies $\tau_r^{sp}(\hat{c})$ if the costs $\hat{c} \in [\hat{c}_a^{sp}, \hat{c}_b^{sp})$ are low.*

Proof Analysis and formal discussion in the text.

5.8 Extensions

5.8.1 Migration vs. education as signalling devices

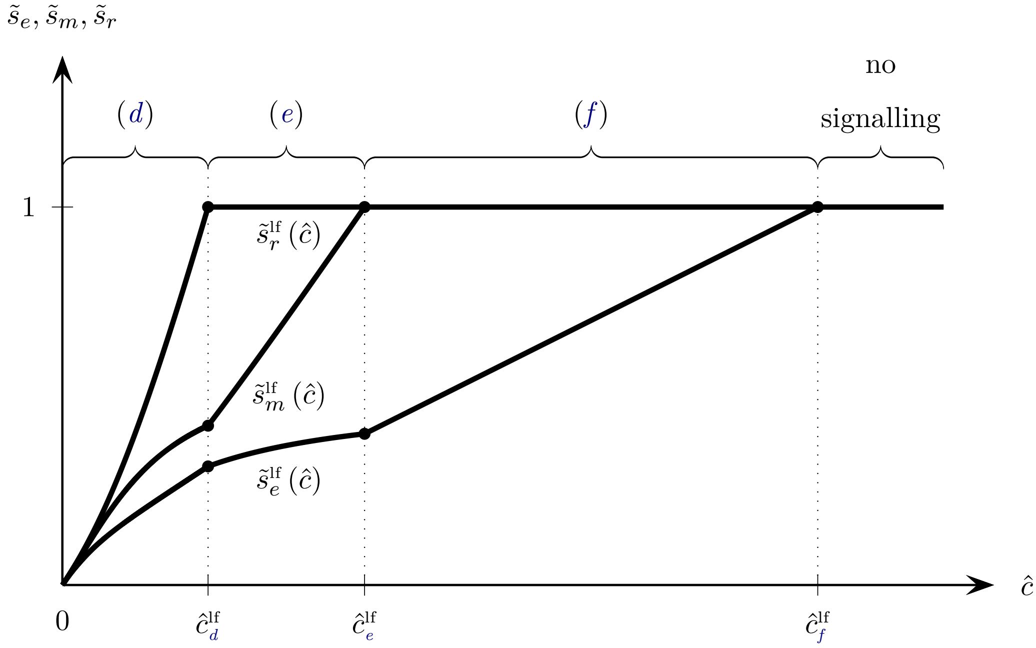
In our baseline model the (repeated) use of costly stays abroad is the only option for workers to signal their otherwise unobservable skills. In reality of course there is a wide range of possible signalling devices with education presumably being the best known example (cf. Spence, 1973). To see whether the existence of alternative signalling devices leads to a crowding out effect, potentially replacing temporary and permanent migration as signals, we extend our baseline model towards a setting with $t = 1, 2, 3$ periods, in which individuals have to decide whether to get educated at age $t = 1$ before temporary or permanently going abroad at age $t = 2, 3$. Education thereby purely acts as a signal and, hence, is assumed to have no human-capital-accumulation effect.[36] To ensure comparability with the baseline model from Section 5.6 we furthermore assume that the (signalling) costs of education and migration are the same $c_e = c_m = c > 0$ and moreover constant over time.[37] Going through the same steps as in the baseline

[36]For models in which the human-capital-accumulation effect of education is instrumental for (temporary) migration decisions see Dustmann and Weiss (2007); Dustmann, Fadlon, and Weiss (2011); Dustmann and Glitz (2011).

[37]In Section 5.4.3 we and show in the static model version that two alternative signals (education vs. migration) may coexists as long as the associated signalling costs $c_e \neq c_m$ are not *too* similar. The simultaneous coexistence

model from Section 5.6 we document the sequential use of education as well as temporary and permanent migration as signalling devices in Appendix A.30. The resulting signalling pattern is then depicted in Figure 5.15. As evident from scenario (f) in Figure 5.15 education may indeed

Figure 5.15: *Alternative signalling devices*



replace temporary and permanent migration as signalling devices if the associated costs for all three signals are sufficiently high, i.e. $\hat{c} \in [\hat{c}_e^{\text{lf}}, \hat{c}_f^{\text{lf}})$ with $\hat{c}_e^{\text{lf}} = 3/5 < \hat{c}_f^{\text{lf}} = 3/2$. However, for the low- and medium-cost scenarios (d) and (e) we find that education, temporary migration and even permanent migration may coexist as signalling devices, provided that the costs for these signals are sufficiently low, i.e. $\hat{c} < \hat{c}_e^{\text{lf}}$ or $\hat{c} < \hat{c}_d^{\text{lf}}$, respectively, with $\hat{c}_d^{\text{lf}} \approx 3/10$. Proposition 5.8.1 summarises the results.

Proposition 5.8.1 *Provided the cost for education and migration are the same, we find that education and temporary migration as well as education, temporary migration, and permanent migration may coexist as signalling devices, if the associated costs are sufficiently small, i.e.* $\hat{c} < \hat{c}_e^{lf}$ *or* $\hat{c} < \hat{c}_d^{lf}$*, respectively.*

---

of both signals thereby is explained by the selection of high-skilled (medium-skilled) workers into the more (less) expensive signal.

Proof Analysis in the text and formal discussion in Appendix A.30.

5.8.2 The strategic effect of initial emigration

In order to highlight the strategic effect of initial emigration decisions on later return decisions we introduce the discounting factor $\delta \in [0,1]$. Going through the same steps as in Section 5.6, thereby discounting future costs and benefits at factor δ (cf. Appendix A.31), we are able to rewrite the emigration cutoff:

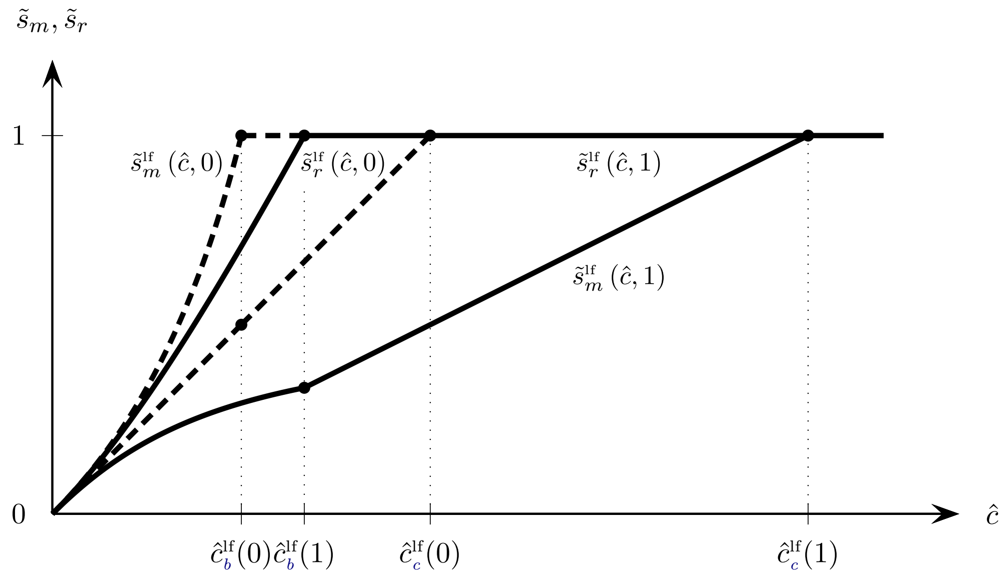
$$\tilde{s}_m^{\text{lf}}(\hat{c},\delta) = \begin{cases} 2\hat{c} & \forall\ \hat{c} \in [0, \hat{c}_a^{\text{lf}}(\delta)) \qquad \Leftrightarrow (a), \\ \dfrac{1 + 2(1+\delta)\hat{c} - \sqrt{1 - 4(1-\delta)\hat{c} + 4(1+\delta)^2\hat{c}^2}}{2} & \forall\ \hat{c} \in [\hat{c}_a^{\text{lf}}(\delta), \hat{c}_b^{\text{lf}}(\delta)) \Leftrightarrow (b), \\ \dfrac{2\hat{c}}{1+\delta} & \forall\ \hat{c} \in [\hat{c}_b^{\text{lf}}(\delta), \hat{c}_c^{\text{lf}}(\delta)) \Leftrightarrow (c), \end{cases} \tag{5.26''}$$

as well as for the return cutoff:

$$\tilde{s}_r^{\text{lf}}(\hat{c},\delta) = \begin{cases} 2\hat{c} & \forall\ \hat{c} \in [0, \hat{c}_a^{\text{lf}}(\delta)) \qquad \Leftrightarrow (a), \\ \dfrac{4\hat{c}}{1 - 2(1+\delta)\hat{c} + \sqrt{1 - 4(1-\delta)\hat{c} + 4(1+\delta)^2\hat{c}^2}} & \forall\ \hat{c} \in [\hat{c}_a^{\text{lf}}(\delta), \hat{c}_b^{\text{lf}}(\delta)) \Leftrightarrow (b), \\ 1 & \forall\ \hat{c} \in [\hat{c}_b^{\text{lf}}(\delta), \hat{c}_c^{\text{lf}}(\delta)) \Leftrightarrow (c), \end{cases} \tag{5.27''}$$

with $\hat{c}_a^{\text{lf}}(\delta) = 0 < \hat{c}_b^{\text{lf}}(\delta) = (1+\delta)/[2(2+\delta)] < \hat{c}_c^{\text{lf}}(\delta) = (1+\delta)/2$. We depict both cutoffs jointly in Figure 5.16 and distinguish between the baseline scenario without discounting, i.e. $\delta = 1$, and a scenario with $\delta = 0$ in which individuals are assumed to be myopic. The latter case thereby serves as a natural benchmark for our baseline model, in which workers emigrate strategically at age $t = 1$, thereby taking into account the later return decisions of their fellow-migrants at age $t = 2$. By raising the discounting factor $\delta \in [0,1]$ gradually from zero to one two effects can be identified: On the one hand, both $\tilde{s}_m^{\text{lf}}(\delta)$ and $\tilde{s}_r^{\text{lf}}(\delta)$ rotate clockwise, while, on the other hand $\tilde{s}_m^{\text{lf}}(\delta)$ ($\tilde{s}_r^{\text{lf}}(\delta)$) becomes more (less) concave (convex) in the range $\hat{c} \in [\hat{c}_a^{\text{lf}}(\delta), \hat{c}_b^{\text{lf}}(\delta))$. The first effect results as workers put more weight on the future benefits from staying abroad, which are discounted at $\delta \in [0,1]$. The adjusted weighting is then what renders temporary migration more attractive, given that the associated costs only accrue at age $t = 1$, while the expected gains continue to pay off at age $t = 2$. The second effect is then what reflects the strategic interaction

Figure 5.16: *The strategic effect of initial emigration*



between initial emigrants ($M \times 1$) at age $t = 1$ and later permanent emigrants ($M \times 2$) at age $t = 2$. Thereby the initial emigrants ($M \times 1$) anticipate at age $t = 1$ that the permanent emigration of the most high-skilled initial emigrants ($M \times 2$) with $s \geq \tilde{s}_r^{\text{lf}}(\delta)$ at age $t = 2$ will depress the wages of *all* returnees ($R \times 2$) including the critical emigrant $\tilde{s}_m^{\text{lf}}(\delta)$. To alleviate these wage losses we find that initial emigration at age $t = 1$ is strategically reduced (more concave shape of $\tilde{s}_m^{\text{lf}}(\delta)$) in order to limit the potential for permanent emigration $1 - \tilde{s}_m^{\text{lf}}(\delta)$ at age $t = 2$ in line with $\tilde{s}_r^{\text{lf}}(\tilde{s}_m^{\text{lf}}) = 2\hat{c}/(1 - \tilde{s}_m^{\text{lf}})$ from Eq. (5.22). As the potential for permanent emigration $1 - \tilde{s}_m^{\text{lf}}(\delta)$ at age $t = 2$ is reduced, less workers will select into permanent emigration at given cost $\hat{c}$, which finally explains the reduced convexity in the shape of $\tilde{s}_r^{\text{lf}}(\delta)$.

5.8.3 Existence of a permanent-migration-only equilibrium

In the Section 5.6 we argue that the permanent-migration-only scenario (*a*) with $0 < \tilde{s}_m = \tilde{s}_r < 1$ cannot exist if the periodical costs $c_t > 0$ of staying abroad are non-decreasing in workers' age $t = 1, 2$. We now extend our baseline model to account exactly for the case in which the costs c_t decline with workers' age $t = 1, 2$. For this purpose let us normalise c_1 to $c_1 = c > 0$, such that c_2 can be expressed as $c_2 = (1 - \gamma)c \geq 0$ with $\gamma \in [0, 1]$ being the percentage decline in costs c as individuals become older. Replacing c_1 and c_2 in Eqs. (5.26′) and (5.27′) by c and $(1-\gamma)c$ then gives us the corresponding emigration and return cutoffs, $\tilde{s}_m^{\text{lf}}(\hat{c},\gamma)$ and $\tilde{s}_r^{\text{lf}}(\hat{c},\gamma)$, which allow us to solve for the cost thresholds $0 \leq \hat{c}_a^{\text{lf}}(\gamma) = \gamma/(2-\gamma)^2 \leq \hat{c}_b^{\text{lf}}(\gamma) = 1/(3-2\gamma) \leq \hat{c}_c^{\text{lf}}(\gamma) = 1$. As we would expect the baseline model, featuring no permanent-migration-only equilibrium, results for $\gamma = 0$, since then $\hat{c}_a^{\text{lf}}(0) = 0$. On the contrary if $\hat{c}_a^{\text{lf}}(\gamma) \in (0,1)$ for $\gamma \in (0,1)$, there always exists a low-cost scenario (a) with $0 < \tilde{s}_m^{\text{lf}}(\hat{c},\gamma) = \tilde{s}_r^{\text{lf}}(\hat{c},\gamma) < 1$, in which *all* initial emigrants with $s \geq \tilde{s}_m^{\text{lf}}(\hat{c},\gamma)$ find it optimal to use a repeated stay abroad (permanent migration) as a signalling device at age $t = 2$. Moreover, since $\hat{c}_a^{\text{lf}\prime}(\gamma), \hat{c}_a^{\text{lf}\prime}(\gamma) > 0$ imply $\lim_{\gamma\to 1} \hat{c}_a^{\text{lf}}(\gamma) = \lim_{\gamma\to 1} \hat{c}_b^{\text{lf}}(\gamma) = \hat{c}_c^{\text{lf}}(\gamma) = 1$, we find that the medium- and high-cost scenarios, (c) and (b), are gradually replaced by the low-cost scenario (a) as the costs $c_2 = (1-\gamma)c$ at age $t = 2$ decline from $c_2 = c$ to $c_2 = 0$. Instrumental for the expansion of scenario (a) is thereby the contraction of scenario (b), in which a separating equilibrium with $0 < \tilde{s}_m^{\text{lf}}(\hat{c},\gamma) < \tilde{s}_r^{\text{lf}}(\hat{c},\gamma) < 1$ can only exist if the signalling costs $c_2 > 0$ at age $t = 2$ are sizeable enough.

5.9 Summary

In this chapter we have developed a model that can explain permanent *and* temporary two-way migration of high-skilled individuals between countries at the same level of economic development. In our model high-skilled individuals use costly migration as a way to signal their true skill level. Support for our theory can be found in the pattern of high-skilled permanent and temporary migration among rather similar countries (like the EU15), which, as we have shown, is characterised by a substantial degree of "two-way-ness".

Our static baseline model is extremely simple, but for this very reason it is transparent as well, and it furthermore lends itself to a comprehensive welfare analysis. We identify a negative externality from migration, resulting from the fact that the marginal migrant ignores the negative effect her migration decision has on expected wages of both natives and migrants. As a consequence, there is too much migration in the laissez-faire equilibrium with positive migration cost, and aggregate welfare is lower than in autarky. This does not mean, however, that *all* migration in our model is socially harmful. We show that, if migration cost is sufficiently low, a social planner would choose strictly positive migration levels. The negative migration externality in this case has to be traded off against the better quality of matches within firms that can be achieved due to the existence of a well-defined high-skill group, comprising the migrants.

The negative migration externality is a fundamental feature of our framework, which survives in more general versions of our static model. The persistence of the negative externality notwithstanding, aggregate gains from migration re-emerge as a possible feature of the laissez-faire equilibrium once our baseline framework is amended by standard features known from other migration models. In particular, once we introduce a second factor that is internationally immobile and a complement to labour in production, aggregate gains from migration exist, provided the income share of this factor is sufficiently high and migration cost is sufficiently low. The welfare gains in this case result from a more efficient domestic allocation of internationally immobile factors of production, notably in the absence of any country asymmetries that would normally be responsible for positive welfare effects of migration.

Finally we show that all the results from our static baseline model carry over to a dynamic two-period setting, in which individuals can choose between temporary and permanent migration. Within this framework costly stays abroad generate an easy-to-verify signal for workers, whose skills would otherwise be unobservable, and it is the *repeated* use of this costly signalling option, which shapes the strategic selection of workers into temporary *and* permanent migration.

Chapter 6

The political economy of high-skilled migration when inequality matters

From 1975 to 2005 the number of high-skilled migrants soared by 2 million tertiary educated individuals (cf. Lowell, 2007).[1] In relative terms, high-skilled migration appears to be even more important as the world-wide average skilled emigration rate of 5.4% in 2000 is nearly six times higher than the emigration rate of unskilled workers (cf. Docquier and Marfouk, 2006). While much has been written about the effects of high-skilled migration on sending countries, the welfare and distribution effects for the host countries have only recently been put on the research agenda (cf. Bougheas and Nelson, 2012, 2013).[2]

This chapter extends the work of Bougheas and Nelson (2012), who analyse the political economy of high-skilled migration between two asymmetric countries in a Ricardian-style general equilibrium model. Unlike in Bougheas and Nelson (2012), individuals display an aversion against (disadvantageous) inequality, as proposed by Fehr and Schmidt (1999).[3] A democratic

[1]For detailed breakdown of migration flows to the six main destination countries from 1975 to 2000 refer to Defoort and Rogers (2008), non-OECD destination countries are covered in Docquier, Özden, Parson, and Artuc (2012).

[2]For the literature on the "brain drain" see Grubel and Scott (1966a), Bhagwati and Rodriguez (1975) and Kwok and Leland (1982). For more recent models emphasizing the possibility of a "brain gain" refer to Mountford (1997), Stark, Helmenstein, and Prskawetz (1997, 1998) and Beine, Docquier, and Rapoport (2001).

[3]For empirical evidence supporting the "relative income" hypothesis see Alpizar, Carlsson, and Johansson-

referendum with respect to the host country's (high-skilled) immigration policy will hence not only depend on whether the median voter benefits from immigration through a higher *absolute* (real) income, but also on the change of median voter's rank (i.e. its *relative* position) in the host country's (real) income distribution. Thereby, the median voter in the host country's labour-intensive sector – as in Bougheas and Nelson (2012) – benefits from the *indirect* terms of trade effect induced by skilled immigration, which at a global scale shifts resources into the host country's more efficient skill-intensive industry. To absorb the resulting expansion in the global production of the skill-intensive good, the (relative) world market price of the skill-intensive good has to decline, and this is what leaves the median voter, who is employed in the host country's labour-intensive sector, better off, both in *absolute* and in *relative* terms. Apart from this indirect terms of trade effect skilled immigration has another more direct effect on individuals, which displays an aversion against (disadvantageous) inequality, as it changes the composition of the host country's (real) income distribution. Since skilled workers from the source country immigrate into the top-ranks of the host country's income distribution, the median voter compares to a group of workers, whose skills and hence incomes are upward biased relative to a situation without migration. The increase in disutility from (disadvantageous) inequality aversion associated with this *direct* composition effect then – of course – must be set against the individual welfare gains from the *indirect* terms of trade effect and it is *a priori* not clear, which of both effects dominates. However, following Bougheas and Nelson (2012) in assuming a uniform distribution of workers skills, it can be shown, that the composition effect is always dominated by the terms of trade effect, such that the median voter prefers an equilibrium with high-skilled migration over an equilibrium without migration, even if preferences are specified to reflect an aversion against (disadvantageous) inequality.

Except for the preference specification, which is borrowed from Fehr and Schmidt (1999), the baseline model resembles the one in Bougheas and Nelson (2012), who rely on a standard two-country-two-sector-model with Ricardian technology differences and heterogeneous workers as proposed by Davidson and Matusz (2006), Davidson, Matusz, and Nelson (2007), Bougheas

Stenman (2005), Ferrer-i Carbonell (2005) and Luttmer (2005). For a more direct test of the preference specification by Fehr and Schmidt (1999) refer to Charness and Grosskopf (2001) as well as to Engelmann and Strobel (2004) who try to validate the theoretical results in experimental settings.

and Riezman (2007) to study various aspects in international trade. The choice of the underlying framework thereby reflects the necessity to analyse migration flows within a general equilibrium model with two countries, that are integrated through trade, instead of relying on a partial equilibrium analysis of a single receiving country in order to arrive at reliable conclusions with regard to the political economy of high-skilled migration between democratically governed partner countries in a globalised world economy. In an equilibrium with costly migration, the most high-skilled workers select into the country guaranteeing a higher reward to skill given its superior technology in the skill-intensive sector, as it would be the case in the Roy-Borjas model (cf. Borjas, 1987). The clustering of skilled workers in the country with the more productive skill-intensive industry reinforces this country's comparative advantage in the production of the skill-intensive good and gives rise to a complementarity between migration and trade. An inevitable consequence of the resulting concentration of skills is a decline in the relative price of the skill-intensive good, which benefits workers employed in the labour-intensive sector and thus makes free migration more likely to be the outcome of a democratic referendum.

This chapter is structured as follows. Section 6.1 presents the baseline model. In Section 6.2 the autarky equilibrium is explored, before turning to the migration equilibrium in Section 6.3. Section 6.4 discusses the political economy equilibrium. Section 6.5 concludes.

6.1 Endowments, preferences, and technology

Imagine two countries, Home and Foreign, where the latter is denoted by an asterisk. Both countries are populated by L and L^* heterogeneous individuals, which are all endowed with one unit of raw labour but differ with respect to their skills s. Individuals skills in both countries thereby are defined over the unit interval $s \in [0, 1]$ and follow the same distribution function $F(s)$ with density function $f(s) = dF(s)/ds$.

As in Bougheas and Nelson (2012) the technology is Ricardian and workers either supply raw labour in the labour-intensive sector (with subindex L) or alternatively make use of their skills, when being employed in the skill-intensive sector (with subindex S). Employing their skills in the skill-intensive sector, workers in Home generate A-times the output of worker in Foreign (with $A > A^* \overset{!}{=} 1$). Evaluated at perfectly competitive prices p_S and p_S^* this technology justifies wages

$$w_S(s) = p_S A s, \qquad \text{and} \qquad w_S^*(s) = p_S^* s, \tag{6.1}$$

which positively depend on the respective worker's skill level s or s^*, respectively. On the contrary, when being employed in the labour-intensive sector, each worker inelastically supplies one unit of raw labour, which is converted into one unit of output. Under perfect competition and evaluated at prices p_L and p_L^* the corresponding wages then follow as

$$w_L = p_L, \qquad \text{and} \qquad w_L^* = p_L^*. \tag{6.2}$$

To sum up, Home has an absolute *and* a comparative advantage in the production of the skill-intensive good.

Individuals in Home and in Foreign have the same, homothetic preferences, which are explicitly specified for Home but of course apply in the same way for Foreign. In particular, the preference specification follows Fehr and Schmidt (1999) and accounts – in addition to the usual (real) income maximisation motive – for a (possible) aversion against an economy-wide unequal distribution of (real) incomes. Individuals thereby compare their own (real) income within a reference group of constant size, which is assumed to be a representative sub-sample of the respective country of residence's total population.[4] Since nominal income $N(s)$ depends (positively) on individual skill s, the indirect utility function for individual $\hat{s}$ can be written as

$$\begin{aligned} V[p_L, p_S, I(\hat{s})] = R[p_L, p_S, N(\hat{s})] \\ - \beta \int_0^1 \max\{R[p_L, p_S, N(s)] - R[p_L, p_S, N(\hat{s})], 0\} f(s)\, ds \\ - \alpha \int_0^1 \max\{R[p_L, p_S, N(\hat{s})] - R[p_L, p_S, N(s)], 0\} f(s)\, ds. \end{aligned} \tag{6.3}$$

As usual, indirect utility depends on the individual real-income level $R[p_L, p_S, I(\hat{s})]$, which is multiplicatively separable into nominal income $N(\hat{s})$ and the inverse of the aggregate price index $1/P(p_L, p_S)$, as long as preferences are homothetic. In addition to their real incomes, individuals care about the disutility resulting from disadvantageous inequality, which is measured as the

[4] As an alternative way to capture aggregate inequality Yitzhaki (1979) and Stark and Taylor (1991) propose a measure of "relative deprivation", which takes into account only disadvantageous inequality. Aggregate relative deprivation thereby also depends on the economy-wide population size.

average real income in excess of individual $\hat{s}$' real income level $R[p_L, p_S, N(\hat{s})]$. The parameter $\beta \geq 0$ thereby reflects the sensitivity towards this type of inequality. Similarly, individuals care about advantageous inequality with $\alpha \gtreqless 0$ again denoting the sensitivity towards this kind of inequality.[5] In the baseline model, $\beta \geq 0$ and $\alpha = 0$ are assumed. Individuals thus show an aversion against disadvantageous inequality, while being neutral with respect to advantageous inequality. Support for this assumption comes from Ferrer-i Carbonell (2005), who finds that individuals mostly compare "up-wards". Similarly, Loewenstein, Thompson, and Bazerman (1989) observe that individuals have a stark aversion against inequality which works to their disadvantage, while the aversion against advantageous inequality seems to be significantly weaker and less robust. Assuming that individuals' nominal income $N(s)$ is non-decreasing in skill s, Eq. (6.3) simplifies to

$$V(p_L, p_S, \hat{s}) = \frac{N(\hat{s}) - \beta\Omega(\hat{s})}{P(p_L, p_S)}, \tag{6.4}$$

with

$$\Omega(\hat{s}) \equiv \int_{\hat{s}}^{1} [N(s) - N(\hat{s})] f(s)\, ds, \tag{6.5}$$

in which $\beta\Omega(\hat{s}) / P(p_L, p_S) \geq 0$ measures the disutility from disadvantageous inequality aversion, which can be small or large, depending on individual $\hat{s}$' rank in Home's income distribution being either high or low.

Notably, with homothetic preferences, individuals' relative consumption choice is not altered through the introduction of inequality aversion. Given their absolute *and* relative position in the income distribution, reflected by $N(\hat{s})$ and $\beta\Omega(\hat{s})$ in Eq. (6.4), individuals still seek to minimise their consumption expenditure, which for the optimal consumption ratio of x_S to x_L equals the aggregate price index $P(p_S, p_L)$ in Eq. (6.4). Optimality thereby requires that the relative goods price p_S/p_L are equated to the marginal rate of substitution between both commodities. Denoting the elasticity of substitution between both goods by $\sigma > 0$, the corresponding *inverse* relative demand function follows as

$$\frac{p_S}{p_L} = \left(\frac{x_S}{x_L}\right)^{-\frac{1}{\sigma}}. \tag{6.6}$$

[5]For an alternative specification of preferences incorporating both, advantageous and disadvantageous inequality aversion, see Ockenfels and Bolton (2000).

6.2 The autarky equilibrium

The derivation of the autarky equilibrium in both economies follows in three steps. At first, the *labour market* equilibrium is described in Lemma in 6.2.1, before the *goods markets* equilibrium is explained in Lemma 6.2.2. Finally, both Lemmas are combined, which gives the solution to the autarky equilibrium.

Labour markets in both countries clear if critical workers $\tilde{s}, \tilde{s}^* \in [0,1]$ exist, who are indifferent between getting employed in the labour- or in the skill-intensive sector. Choosing the labour-intensive good as *numéraire* (i.e. $p_L \stackrel{!}{=} 1$ such that $p = p_S/1$), the corresponding indifference conditions can be summarised in Lemma 6.2.1, which follow directly from the comparison of indirect utilities in Eq. (6.4) evaluated at the sectoral wage rates from Eqs. (6.1) and (6.2), respectively.

Lemma 6.2.1 *The indifferent workers $\tilde{s}$ and $\tilde{s}^*$ are implicitly defined by*

$$V(p, w_L) = V[p, w_S(\tilde{s})] \quad \Leftrightarrow \quad p = \frac{1}{A\tilde{s}}, \tag{6.7}$$

$$V^*(p^*, w_L^*) = V^*[p^*, w_S^*(\tilde{s}^*)] \quad \Leftrightarrow \quad p^* = \frac{1}{\tilde{s}^*}. \tag{6.8}$$

Proof Analysis and formal discussion in the text. ■

Intuitively, all workers with skills below $\tilde{s}$ and $\tilde{s}^*$, respectively, are employed in the labour-intensive sector, while all workers with skills above these thresholds are employed in the skill-intensive sector. Thereby, the allocation of workers is driven by *both* the workers' absolute *and* relative position in the country's overall income distribution. High-skilled workers with $s > \tilde{s}$ ($s^* > \tilde{s}^*$) not only have higher absolute incomes, they also accompany a higher rank in the overall income distribution, which, in addition to the utility gain from higher absolute consumption, reduces the disutility from workers' inequality aversion. Consequently, if relative prices change in favour of the skill (labour)-intensive good, the cutoffs $\tilde{s}$ and $\tilde{s}^*$ fall (rise), which leads to an expansion of the skill (labour)-intensive sector.

Goods markets clearing in Home and Foreign requires the equality of relative demand and relative supply. While relative demand (for Home) is given by Eq. (6.6), relative supply follows directly from the sorting of workers into sectors, implying sectoral supplies $y_L = F(\tilde{s})L$ as well as $y_S = A\int_{\tilde{s}}^{1} sf(s)dsL$ for Home and $y_L^* = F(\tilde{s}^*)ds^*L^*$ as well as $y_S^* = \int_{\tilde{s}^*}^{1} s^*f(s^*)ds^*L^*$ for Foreign. Equating relative demands and supplies in both countries finally yields the equilibrium conditions for the goods markets, which are summarised in Lemma 6.2.2.

Lemma 6.2.2 *Equilibrium conditions for the goods markets are given by*

$$p = \left(\frac{y_S}{y_L}\right)^{-\frac{1}{\sigma}} \quad \Leftrightarrow \quad p = g\left(\tilde{s}\right) \equiv \left(\frac{F\left(\tilde{s}\right)}{A\int_{\tilde{s}}^{1} sf\left(s\right)ds}\right)^{\frac{1}{\sigma}}, \tag{6.9}$$

$$p^* = \left(\frac{y_S^*}{y_L^*}\right)^{-\frac{1}{\sigma}} \quad \Leftrightarrow \quad p^* = g^*\left(\tilde{s}^*\right) \equiv \left(\frac{F\left(\tilde{s}^*\right)}{\int_{\tilde{s}^*}^{1} s^*f\left(s^*\right)ds^*}\right)^{\frac{1}{\sigma}}, \tag{6.10}$$

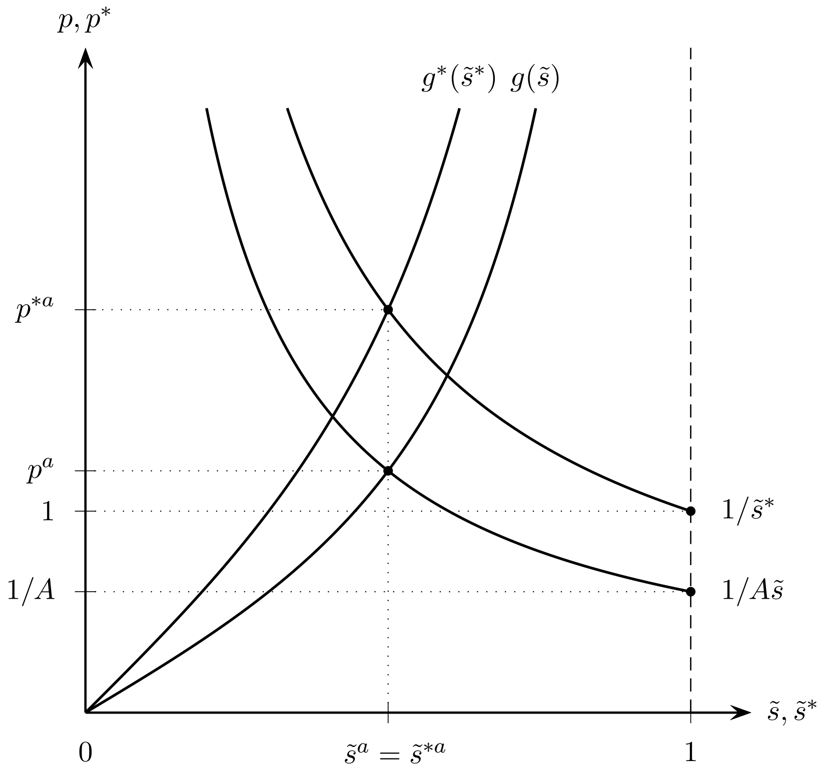
with $\partial g\left(\tilde{s}\right)/\partial\tilde{s} > 0$ *and* $\partial g^*\left(\tilde{s}^*\right)/\partial\tilde{s}^* > 0$.

Proof Analysis and formal discussion in the text. ■

According to Lemma 6.2.2 the (relative) prices of the skill-intensive good p in Home and p^* in Foreign depend on the relative size of the skill-intensive sectors in the two economies, which is captured by the cutoffs $\tilde{s}, \tilde{s}^* \in [0,1]$, respectively. Thus, if for example the labour-intensive sector in Home expands, because the cutoff $\tilde{s}$ has increased, this will shift relative supply y_S/y_L towards the labour-intensive good. To absorb the excess supply of the labour-intensive good, the (relative) price for the skill-intensive good p has to go up and this is why $g(\tilde{s})$ is upward sloping in $\tilde{s}$.

The autarky equilibrium can now be graphically solved in Figure 6.1 by combining the conditions for the *labour market* and *goods market* equilibrium from the Lemmas 6.2.1 and 6.2.2. In equilibrium, the cutoffs $\tilde{s}, \tilde{s}^* \in [0,1]$ along with the corresponding (relative) prices p and p^* are uniquely pinned down through the intersection points of the two equilibrium conditions in Home and Foreign. The *labour market* equilibrium in Foreign thereby at each (relative) price p^* has employed less workers in the skill-intensive sector, which, due to Foreign's less efficient production technology, pays lower wages than the comparable sector in Home. Nevertheless, a factor allocation characterised by $\tilde{s}^* > \tilde{s}$ should not be taken for granted, given that the *goods market* equilibrium in Foreign, at the same time, supports a higher (relative) price for the skill-intensive good p^*, for each allocation of factors implied by the cutoff $\tilde{s}^*$. The price in Foreign is higher, since compared to Home, less skill-intensive output is produced at a given allocation of

Figure 6.1: *The autarky equilibrium in Home and Foreign*



factors, such that the skill-intensive good is relatively scarce in Foreign. Both effects (partially) offset each other and it is hence a priori not clear, in which country (relatively) more workers are employed in the labour- or the skill-intensive sector. Figure 6.1 depicts the knife-edge case with $\tilde{s} = \tilde{s}^*$, which results for $\sigma = 1$ and resembles the familiar result that sector-specific Hicks-neutral technological change has no effect on the allocation of workers and on relative factor prices with Cobb-Douglas preferences (cf. Krugman, 2000; Xu, 2001).

6.3 The migration equilibrium

The migration equilibrium again is derived in three steps. At first Lemma 6.3.1 addresses the selection into migration. Goods markets are assumed to clear at a global scale and the corresponding *goods market* equilibrium for the integrated world economy is derived in Lemma

6.3.2. Finally, the *goods market* equilibrium condition from Lemma 6.3.2 and the (unchanged) *labour market* equilibrium condition from Lemma 6.2.1 are combined and jointly determine the migration equilibrium and the terms of trade between Home and Foreign.

Migration is costly and the cost of migration are paid in units of the final goods, whose relative spending ratio is assumed to be the same as the relative consumption ratio implied by Eq. (6.6). As a consequence, the resource use associated with migration ($\gamma \geq 0$) is evaluated at the aggregate price index $P^W(p^W, 1)$ and individual migration decisions do not depend on the terms of trade. Thus, if international migration and the associated reallocation of labour across sectors in Home and Foreign feeds back into the relative world market price, this does not distort individual migration decisions.

From the perspective of a skilled foreign worker with $s^* > \tilde{s}$ migration is beneficial if the access to Home's superior production technology ($A > A^* = 1$) compensates for a comparatively lower rank in Home's income distribution and moreover covers the migration cost $\gamma > 0$. Formally, a threshold-level $\breve{s}^* \in [\tilde{s}, 1]$ can be defined such that all workers with $s^* \geq \breve{s}^*$ prefer to work in Home, while workers with $s^* < \breve{s}^*$ prefer to stay in Foreign. The critical worker $\breve{s}^* \in [\tilde{s}, 1]$ thereby trades off a higher absolute income, due to the access to Home's superior technology, against the top rank in Foreign's income distribution, which is associated with zero disutility from disadvantageous inequality aversion. High skills in this trade off not only ensure a higher absolute income but also a higher rank in Home's income distribution, which renders migration to Home – ceteris paribus – more attractive. The corresponding indifference condition then follows from $V[p^W, w_S(\breve{s}^*)] - P(p^W, 1)\gamma = V^*[p^W, w^*(\breve{s}^*)]$, which evaluated at wages $w_S(\breve{s}^*) = p^W A\breve{s}^* \geq w^*(\breve{s}^*) = \max\{w_S^*(\breve{s}^*), w_L^*\}$ with $w_S^*(\breve{s}^*) = p^W \breve{s}^*$ and $w_L^* = 1$ simplifies to

$$A\left[\breve{s}^* - \beta\Omega\left(\breve{s}^*\right)\right] - \gamma = \max\left\{\breve{s}^*, \tilde{s}^*\right\}, \tag{6.11}$$

in which $\gamma < A - 1$ is assumed to ensure that a migration equilibrium with $\breve{s}^* < 1$ exists.

For skilled workers in Home with $s > \tilde{s}$ migration is considerably less attractive, which has two reasons. At first, these workers already start with Home's superior technology ($A > 1$), which they have to give up when migrating abroad. And secondly, migration as such is costly ($\gamma > 0$). The corresponding indifference condition follows from the comparison of $V[p^W, w_S(\breve{s})] = V^*[p^W, w_S^*(\breve{s})] - P(p^W, 1)\gamma$, which evaluated at wages $w_S(\breve{s}) = p^W A\breve{s} \geq w_S^*(\breve{s}) = p^W\breve{s}$ simplifies to

$$A\left[\breve{s}-\beta\Omega\left(\breve{s}\right)\right]=\breve{s}-\gamma, \tag{6.12}$$

in which $\breve{s}$ assumes the role of the skill threshold above (below) which workers stay in Home (migrate to Foreign). Comparing Eqs. (6.11) and (6.12), it can be safely concluded that workers from Home are relatively more inclined to live in Home (i.e. $\breve{s} < \breve{s}^*$) whenever $\gamma > 0$. A sufficient condition to rule out migration from Home to Foreign is $\gamma > A\beta\Omega\left(0\right)$, which states that the disutility from disadvantageous inequality aversion in Home should not exceed the cost of migration for any worker. Lemma 6.3.1 summarizes:

Lemma 6.3.1 *For migration cost $A\beta\Omega\left(0\right) < \gamma < A-1$ workers with $s^* \geq \breve{s}^*$ migrate to Home, while workers with $s^* < \breve{s}^*$ stay in Foreign. Thereby the cutoff $\breve{s}^* \in [\tilde{s}^*, 1]$ is decreasing in A and increasing in β and γ.*

Proof Analysis and formal discussion in the text. ■

Goods markets clear at a global scale. Hence, the relative world market supply y_S^W/y_L^W with $y_S^W = y_S + y_S^*$ and $y_L^W = y_L + y_L^*$ has to equal the relative world market demand in Eq. (6.6) evaluated at the (relative) world market price p^W. Thereby, it can be inferred from Lemma 6.2.1, that, at a common (relative) world market price $p^W = p = p^*$, Home – due to its superior technology – employs (relatively) more workers in its skill-intensive sector ($A\tilde{s} = \tilde{s}^*$ implies $\tilde{s} < \tilde{s}^*$). Hence for a given migration cutoff $\breve{s}^* \in [\tilde{s}, 1]$ two possible patterns of specialisation exist. Suppose $\tilde{s} < \tilde{s}^* < \breve{s}^*$. Then both countries are incompletely specialised and produce both goods. Thereby the employment of domestic workers in Home is divided across sectors. Workers with skills $s \geq \tilde{s}$ are employed in the skill-intensive industry and workers with skills $s < \tilde{s}$ end up in the labour-intensive industry. On the contrary, migrants in Home with skills $s^* \geq \breve{s}^*$ are solely employed in Home's skill-intensive industry. In Foreign the remaining non-migrants find employment in both industries, such that all workers with skills $s^* < \tilde{s}^*$ are employed in the labour-intensive sector, while those workers with skills $\tilde{s}^* \leq s^* < \breve{s}^*$ are employed in the skill-intensive sector. Now suppose $\tilde{s} < \breve{s}^* \leq \tilde{s}^*$. Then only Home produces both goods and Foreign specialises in the production of the skill-intensive good. While the sorting of workers according to the thresholds $\tilde{s}$ and $\breve{s}^*$ in Home is the same as before, in Foreign the migration cutoff $\breve{s}^*$ replaces the industry cutoff $\tilde{s}^*$ and all workers with skills $s < \breve{s}^*$ are employed in the labour-intensive industry. Aggregating up the sectoral outputs and replacing $x_S^W/x_L^W = y_S^W/y_L^W$ in Eq. (6.6) then yields the global goods market equilibrium, which is summarised in Lemma 6.3.2.

Lemma 6.3.2 *The global goods market clearing condition is given by*

$$p^W = \left(\frac{y_S^W}{y_L^W}\right)^{-\frac{1}{\sigma}} \quad \Leftrightarrow \quad p^W = g^W\left(\tilde{s}, \tilde{s}^*, \breve{s}^*\right), \tag{6.13}$$

where

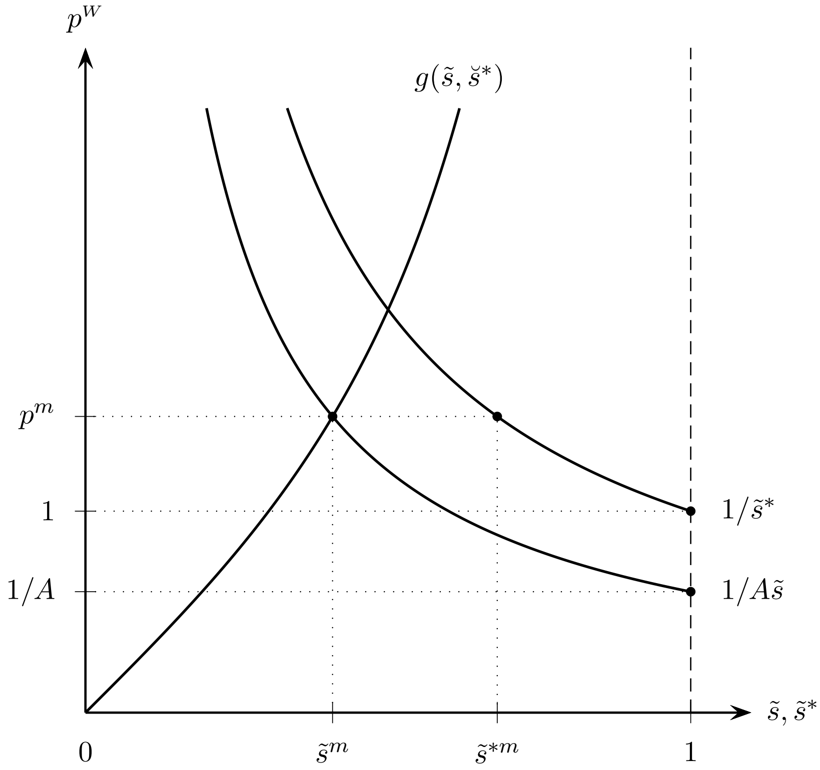
$$g^W\left(\tilde{s}, \tilde{s}^*, \breve{s}^*\right) \equiv \begin{cases} \left(\frac{F(\tilde{s})L + F(\tilde{s}^*)L^*}{A\int_{\tilde{s}}^1 sf(s)dsL + A\int_{\breve{s}^*}^1 s^*f(s^*)ds^*L^* + \int_{\tilde{s}^*}^{\breve{s}^*} s^*f(s^*)ds^*L^*}\right)^{1/\sigma} & \text{if } \tilde{s} \le \tilde{s}^* \le \breve{s}^*, \\ \left(\frac{F(\tilde{s})L + F(\breve{s}^*)L^*}{A\int_{\tilde{s}}^1 sf(s)dsL + A\int_{\breve{s}^*}^1 s^*f(s^*)ds^*L^*}\right)^{1/\sigma} & \text{if } \tilde{s} \le \breve{s}^* < \tilde{s}^*. \end{cases} \tag{6.14}$$

Proof Analysis and formal discussion in the text. ■

Replacing $\tilde{s}^*$ in Eq. (6.14) by $\tilde{s}^* = A\tilde{s}$, which follows from Lemma 6.2.1 for $p^W = p = p^*$, reveals that p^W and $\tilde{s}$ are positively linked through the global goods market equilibrium. The intuition between this link is the following: A parallel rise in $\tilde{s}$ and, hence, in $\tilde{s}^*$ shifts resources from the skill to the labour-intensive sector and, thus, expands the relative supply of the labour-intensive good, which can only be absorbed by the market if at the same time the consumption of the skill-intensive good becomes less attractive, which is the case when the relative world market price p^W goes up.

Given that $\breve{s}^*$ in Eq. (6.13) according to Eq. (6.11) only depends on exogenous parameters, the goods market equilibrium condition in Eq. (6.13) can be plotted as an upward sloping curve $g(\tilde{s}, \breve{s}^*)$ in Figure 6.2, whereas $g(0, \breve{s}^*) = 0$, $\partial g(\tilde{s}, \breve{s}^*)/\partial \tilde{s} > 0$ as well as $g(1, \breve{s}^*) \to \infty$ follow from Eq. (6.14), when replacing $\tilde{s}^*$ by $\tilde{s}^* = A\tilde{s}$. In order to find the equilibrium values of p^W and $\tilde{s}$, the goods market equilibrium condition $g(\tilde{s}, \breve{s}^*)$ can be combined with Home's labour market equilibrium condition in Eq. (6.7), which is depicted as downward sloping curve in Figure 6.2. The intersection point of both curves then determines the equilibrium values of p^W and $\tilde{s}$. Finally, the cutoff $\tilde{s}^*$ can be read off from Foreign's labour market equilibrium condition $p^W = 1/\tilde{s}^*$ once p^W has been determined.

Figure 6.2: *The migration equilibrium*



To explore how migration alters the allocation of factors (i.e. the cutoffs $\tilde{s}$ and $\tilde{s}^*$) as well as the terms of trade p^W, consider a reduction in migration cost $\gamma > 0$, which is reflected by a decline in the migration cutoff $\breve{s}^*$. From the inspection of Eq. (6.14) it follows that the influx of migrants is associated with a clockwise rotation of goods market equilibrium condition $g(\tilde{s}, \breve{s})$ in Figure 6.2. As an immediate consequence the relative world market price p^W declines, while the industry cutoffs $\tilde{s}$ and $\tilde{s}^*$ go up. Interestingly, the intuition for these adjustments depend on the underlying specialisation patterns in both economies. If both countries are incompletely specialised (i.e. $\tilde{s} \leq \tilde{s}^* \leq \breve{s}^*$), worker migrate between countries but within the skill-intensive sector. At a global scale then more workers are employed at Home's superior technology, which – at a notionally fixed world market price – leads to an excess supply of the skill-intensive good and, hence, to a downward shift in the goods market equilibrium condition $g(\tilde{s}, \breve{s}^*)$. For the world market to rebalance, the relative price of the skill-intensive good p^W has to decline, which

– *ceteris paribus* – makes employment in both countries' skill-intensive industries less attractive and, hence, results in an increase in $\tilde{s}$ and $\tilde{s}^*$. For $\tilde{s} \leq \breve{s}^* < \tilde{s}^*$ a different mechanism is in place. Since Foreign is specialised in the production of the skill-intensive good, workers migrate from Foreign's labour-intensive sector into Home's skill-intensive sector. Thus, at a global scale migration is associated with a shift of resources from labour to skill-intensive production. As before the skill-intensive production expands globally, which shifts the goods market equilibrium condition downwards and hence results in a smaller relative world market price p^W and higher cutoffs $\tilde{s}$ and $\tilde{s}^*$. Summing up we can conclude:

Proposition 6.3.3 *Migration from Foreign to Home is associated with a decline of the relative world market price and a reduction in the employment of domestic workers in the production of the skill-intensive good.*

Proof Analysis and formal discussion in the text. ■

Note that the resulting migration pattern implicitly determines the trade flows between Home and Foreign. If a larger fraction of workers emigrate from Foreign, disproportionally more workers are employed at Home's advanced technology in the skill-intensive sector. As a consequence the relative supply of the skill-intensive good from Foreign falls short of the corresponding relative supply from Home. With identical preferences, the free trade equilibrium then features skill-intensive exports from Home, and labour-intensive exports from Foreign.[6].

6.4 The political economy

With Proposition 6.3.3 at hand the political economy of high-skilled migration can be explored by focussing on the impact that high-skilled migration from Foreign to Home has on the median voter in both countries. Thereby, it is assumed that the median voter is low-skilled and hence in an equilibrium without migration employed in the labour-intensive sector. Moreover, migrants, which by construction always benefit from migration, are excluded from elections to make sure

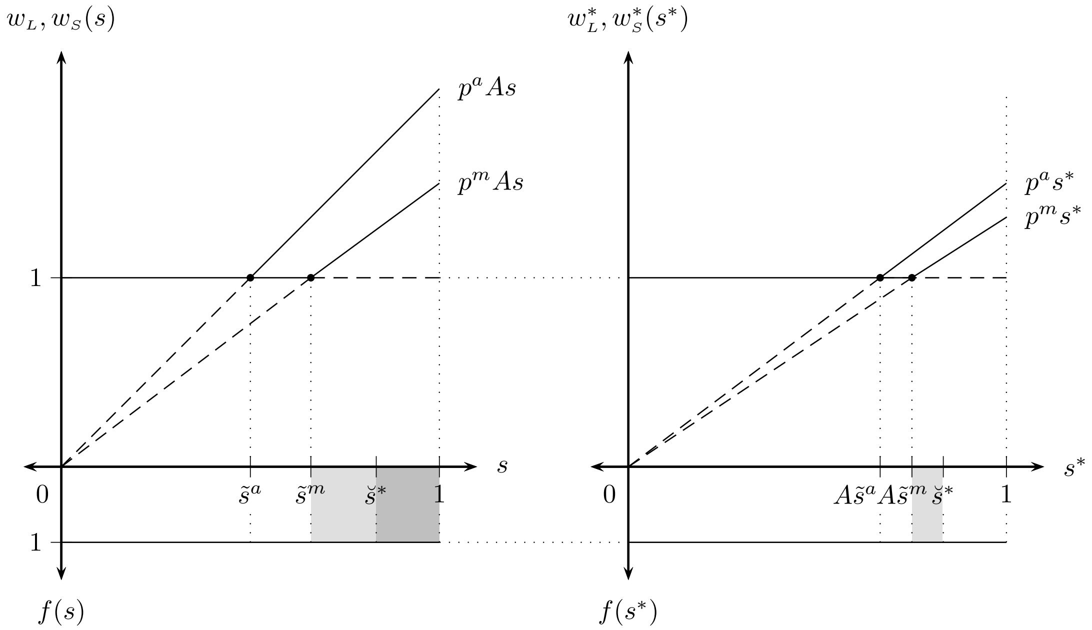
[6] For a detailed analysis of the complementarity between migration and trade in the underlying modeling framework compare to Bougheas and Nelson (2012)

that results are not driven by migration induced composition effects in the destination country's labour force.

Beginning the analysis with the median voter in Foreign, it is obvious that any migration equilibrium is supported by a majority of Foreign's population. The median voter thereby prefers an equilibrium with migration over an equilibrium without migration for two reasons. On the one hand, such an equilibrium supports a lower relative world market price for the skill-intensive good, which is equivalent to a real income gain for low-skilled workers. On the other hand, the equilibrium is characterised by the outmigration of the most high-skilled workers from Foreign and hence features less disadvantageous inequality for those workers left behind. Figure 6.3 (right panel) illustrates both effects. As predicted by Proposition 6.3.3, the (relative) price of the skill-intensive good $p^m < p^a$ declines, which is reflected by a downward rotation of the wage function from Eq. (6.1) in Figure 6.3. The associated real income gain (loss) for workers in the labour (skill)-intensive sector renders employment in the labour (skill)-intensive sector relatively more (less) attractive, such that some domestic workers leave the skill-intensive sector to find employment in the labour-intensive sector. As the industry cutoff $\tilde{s}^*$ goes up, there are less workers with skill $s^* \geq \tilde{s}^*$, who earn a wage rate $w_S^*(s^*) \geq w_L^* = 1$. At the same time, workers with skill $s^* \geq \breve{s}^*$ migrate abroad, which thins out the income distribution from the top and only leaves workers with skills $\tilde{s}^* \leq s^* < \breve{s}^*$ (light grey area) in Foreign's skill-intensive industry. Taking stock, for Foreign's median voter the disutility from disadvantageous inequality is reduced in two ways. The direct outmigration of skilled workers from the top of Foreigns income distribution directly reduces disadvantageous inequality, while the migration-induced terms of trade effect indirectly raises (lowers) the real income of workers in the labour (skill)-intensive sector, which further reduces the disutility from disadvantageous inequality.

Home's median voter may benefit or lose from migration. Similar to the median voter in Foreign, the median voter in Home thereby welcomes the induced terms of trade effect (i.e. $p^m < p^a$) that is associated with high-skilled migration from Foreign to Home. As a consequence of this price change the wage rate $w_S^*(s^*)$ from Eq. (6.1) in Figure 6.3 declines, which translates into a real income gain (loss) for workers in the labour (skill)-intensive sector. On top of these (real) income gains the median voter also benefits from the (real) income losses of the better

Figure 6.3: *The political economy of high-skilled migration*



paid workers in the skill-intensive sector, whose wage premium $w_S(s) \geq w_L$ shrinks in relation to the wage $w_L = 1$ in the labour-intensive sector and hence is associated with less disutility from disadvantageous inequality aversion. Negative consequences for the median voter may arise from the direct effect of migration, i.e. the influx of high-skilled Foreign workers, who find themselves at the top of Home's income distribution after migrating to Home. This inflow of high-skilled workers not only renders Home's income distribution less equal, but also affects the median voter, who now compares to a group of workers whose skills *and* wages are upward biased compared to an equilibrium without migration. To highlight this composition effect, the mass of domestic workers with skills $\tilde{s} \leq s < \breve{s}^*$ is depicted in the left panel of Figure 6.3 in light grey, while the masses of domestic *and* foreign workers with skills $\breve{s}^* \leq s \leq 1$ and $\breve{s}^* \leq s^* \leq 1$, respectively, are depicted in dark grey. Thus, it is *a priori* not clear whether there are net gains for median voter from the migration induced terms of trade effect, or net losses due to increased disutility from disadvantageous inequality aversion.

One way to resolve this ambiguity is to assume a specific functional form for the distribution of workers' skills in both economies. Hence – as in Bougheas and Nelson (2012) – workers' skills are assumed to follow an uniform distribution over the unit interval $s, s^* \in [0,1]$. The median voter then always prefers an equilibrium with migration over an equilibrium without migration. Proposition 6.4.1 summarizes the result:

Proposition 6.4.1 *If skills are uniformly distributed, i.e.* $f(s) = f(s^*) = 1$*, the median voter prefers any migration equilibrium with* $\breve{s}^* \in [\tilde{s}, 1)$ *over an equilibrium without migration.*

Proof See Appendix A.32.

Another way to ensure that Home's median voter benefits from migration is to consider a migration scenario in which the aforementioned composition effect does not exist. For this to happen, Foreign's migration cutoff $\breve{s}^*$ and Home's industry cutoff $\tilde{s}$ must coincide. Only then the skill bias, that would otherwise result from the positive selection of Foreign migrants for $\breve{s}^* > \tilde{s}$, does not exist. The corresponding migration scenario with $\breve{s}^* = \tilde{s}$ emerges for $\gamma = 0$ in combination with $A[\tilde{s} - \beta\Omega(\tilde{s})] > \tilde{s}$, because only if migration costs are absent and Home's technology is advanced enough to outweigh the utility loss from disadvantageous inequality for all Foreign workers with skill $s^* \geq \tilde{s}$, these workers will migrate to Home and nobody find it optimal to migrate from Home to Foreign. While – admittedly – this is a special case, it nevertheless directs our attention to the fact that the migration induced composition effect has a non-monotonic nature, being strong for intermediate levels of migration and weak for very high ($\breve{s}^* \to \tilde{s}$) or very low ($\breve{s}^* \to 1$) migration flows. Holding the induced terms of trade effect fixed, this implies that high-skilled migration is more likely to be welcome by Home's median voter if it has a moderate effect on the skill composition of Home's skill-intensive sector.

6.5 Summary

This chapter develops a theory of high-skilled migration with individuals that display an aversion against disadvantageous inequality. Thereby, individuals not only maximise the absolute of their (real) income, but are also concerned about their relative position within the economy-wide (real) income distribution. The median voter in the destination country is affected through high-skilled immigration in two ways: As in Bougheas and Nelson (2012), there are absolute income gains as high-skilled immigration changes the terms of trade to the advantage of the median voter, who is employed in the destination country's labour-intensive industry. However, since the median voter is also concerned about its rank in the destination country's (real) income distribution, there is a second potentially offsetting effect: skilled immigration into the top of the destination country's income distribution increases the disutility from disadvantageous inequality aversion and hence works to the disadvantage of the median voter in the destination country. Taking stock, individuals, which display an aversion against disadvantageous inequality are less likely to benefit and hence to support high-skilled immigration, which could be an explanation why many advanced countries still hesitate to allow for large scale skilled immigration.

Chapter 7

Conclusion

This thesis takes up a bipolar stance on the international integration of national labour markets focussing, on the one hand, on offshoring, i.e. the relocation of production steps towards low-cost locations abroad, and, on the other hand, on labour migration, and in particular the migration of high-skilled workers. Both phenomena are of particular importance in context of the recent wave of globalisation, which unlike previous internationalisation periods, predominantly characterised by the integration of (final) goods markets through international trade, directly impacts previously (more or less) isolated national labour markets. The thesis thereby explores four broad topics: The first two devoted to the analysis of offshoring and the remaining two referring to the analysis of high-skilled migration.

The first topic deals with the welfare and distributional consequences of offshoring when firms are heterogeneous in terms of their productivity. Acknowledging that firms are positively selected with regard to different internationalisation strategies (such as international trade or horizontal foreign direct investments), it is explored how the selection of firms into offshoring affects the outcomes for individual firms, represented through an entrepreneur and its employees, as well as the outcomes for the aggregate economy. Thereby, it turns out that adjustments along the extensive margin of offshoring (between multinational and domestic firms) give rise to novel effects that can not be analysed within traditional neoclassical trade models. Particularly useful in this context is the possibility to differentiate between what happens at the micro-level within single firms and at the macro-level within the aggregate economy, since often, what happens in the aggregate is quite the opposite of what one would expect when guided by firm-level results. For example, it may be well possible that, while offshoring firms become more productive through the access to cheap labour from abroad, average (industry or economy-wide) productivity declines at the same time. Instrumental for such diverging micro- and macro-level effects is the reallocation of employment shares between domestic and offshoring firms, which are differently affected by possibility to source cheap inputs from abroad. In this way, the analysis not only provides a so far missing link between empirically well observed firm-level results and the outcomes for the aggregate economy, but also may serve as a promissing point of departure for further empirical analysis.

The second topic of this thesis takes a closer look on how workers (instead of firms) may react on a given offshoring shock, when they have the possibility to upgrade their skills through investing in costly on-the-job training. In general equilibrium, where all rents from offshoring unavoidably incur to workers, offshoring creates previously unexploited skill-upgrading possibilities by raising workers wages beyond the cost of skill acquisition. The resulting positive link between increased offshoring and individual propensity of skill upgrading is tested using survey data from the German manufacturing. Thereby, it turns out that industry-level offshoring growth rates indeed correlate with the individual probability of on-the-job training in a positive and highly significant way. In obtaining this result, a wide set of individual and firm related characteristics such as technological change at the workplace among other major determinants of individual on-the-job training are taken into account.

Starting out from the empirical observation of remarkably balanced bilateral (temporary and permanent) high-skilled migration stocks, the third topic of my thesis explores the determinants and consequences of two-way high-skilled migration between similar countries. In absence of natural migration incentives (i.e. cross-country asymmetries) the selection of high-skilled workers into costly (temporary and permanent) migration serves as a signalling device, which enables workers to reveal their high but otherwise unobservable skills to future employers abroad. As shown in several extensions, the signalling theory of two-way migration is able to replicate different (observable) stylised fact of international (temporary and permanent) migration. Moreover, it gives rise to some interesting welfare effects. A negative externality from migration, resulting from the fact that the marginal migrant ignores the negative effect her migration decision has on expected wages of both natives and migrants is identified. As a consequence, there is too much migration in the laissez-faire equilibrium with positive migration cost, and aggregate welfare is lower than in autarky. This does not mean, however, that all migration in our model is socially harmful. We show that, if migration cost is sufficiently low, a social planner would choose strictly positive migration levels. The negative migration externality in this case has to be traded off against the better quality of matches within firms that can be achieved due to the existence of a well-defined high-skill group, comprising the migrants.

The fourth topic analyses the political economy of high-skilled migration, when individuals not only care about their absolute income but are also concerned about their (relative) position within the income distribution of their reference group. The median voter in the immigration country then is affected through high-skilled immigration in two ways: On the one hand, there are absolute income gains as high-skilled immigration changes the terms of trade to the advantage of the (rather) low-skilled median voter, who is employed in the destination country's labour-intensive industry. On the other hand, since the median voter is also concerned about its rank in the immigration country's income distribution, there is a second potentially offsetting effect as skilled immigration into the top of the immigration country's income distribution increases the disutility from disadvantageous inequality aversion and hence works to the disadvantage of the median voter in the destination country. Taking stock, individuals, which display an aversion against disadvantageous inequality are less likely to benefit and hence to support high-skilled immigration, which could be an explanation why many advanced countries still hesitate to allow for large scale skilled immigration.

Taken together, this thesis contributes to a ongoing debate on the pros and cons of two highly relevant forms on international labour market integration: offshoring and high-skilled labour migration. Thereby, the analysis is undertaken in form of four different modelling frameworks, tailored to replicate and explain up-to-date empirical observations on the respective topics. The obtained insights together with the tractability of the underlying frameworks may provide a fruitful guidance to the rapidly growing empirical literature on both topics, being, at the same time, a useful point of departure for further theoretical work in these fields.

Chapter 8

Appendix

A Theory appendix

A.1 A continuum of tasks that differ in offshorability

In this extension, we shed light on the firm-internal margin of offshoring, by considering a continuum of tasks that differ in offshorability, as suggested by Acemoglu and Autor (2011). For this purpose, we replace our production function for intermediates in Eq. (3.3) by

$$q(v) = \varphi(v) \exp \int_0^1 \ln \ell(v, \tilde{\eta}) d\tilde{\eta}, \tag{A.1}$$

in which $\ell(v, \tilde{\eta})$ is the input of task $\tilde{\eta} \in [0, 1]$ in the production of $q(v)$. Tasks are symmetric in the labour input they require to be performed and, as in the main text, we impose the additional assumption that one unit of labour must be employed to produce one unit of task $\tilde{\eta}$. However, as in Grossman and Ross-Hansberg (2008), tasks differ in their offshorability and this is captured by an iceberg cost parameter t that is task specific: $t(\tilde{\eta})$. An intuitive way to interpret parameter t is to think of it as task-specific trade cost parameter, implying that total costs of shipping the output of a task $\tilde{\eta}$, whose production has been moved offshore, back to the source country amounts to $t(\tilde{\eta})\tau > 1$. To facilitate the analysis, we impose the additional assumption that $t(1) = 1$, $t(0) = \infty$ and $t'(\tilde{\eta}) < 0$. This implies that tasks are ranked according to their offshorability and it allows us to identify a unique firm-specific $\eta(v)$, which separates the tasks performed at home, $\tilde{\eta} < \eta(v)$, from the tasks performed abroad $\tilde{\eta} \geq \eta(v)$.

Once a firm has decided to engage in offshoring, it is left with two further decisions on how to organise its production, which are taken in two consecutive stages. In stage one, the firm chooses how many tasks to move offshore and sets $\eta(v)$ accordingly, while in stage two, the firm chooses optimal employment in domestic and offshored tasks. As it is common practice, we solve this two stage problem through backward induction and first determine the profit-maximising employment levels for a given $\eta(v)$. For this purpose, we can recollect from the main text that wages paid to domestic and foreign workers are given $w^o(v)$ and $w^*(v)$, respectively. We can write labour demand for domestic and foreign task production as follows: $l^n(v) = \int_0^{\eta(v)} \ell^n(v,\tilde{\eta})d\tilde{\eta} = \eta(v)\ell^n(v)$ and $l^r(v) = \int_{\eta(v)}^1 t\left(\tilde{\eta}\right)\ell^r\left(v,\tilde{\eta}\right)d\tilde{\eta} = \int_{\eta(v)}^1 t\left(\tilde{\eta}\right)d\tilde{\eta}\ell^r\left(v\right)$.[1] Therefore, firm v's cost minimisation problem can be expressed as follows:

$$\min_{l^n(v),\, l^r(v)} \omega^n\left(v\right)l^n\left(v\right) + \omega^r(v)l^r\left(v\right) \quad \text{s.t.} \quad 1 = \varphi\epsilon[\eta\left(v\right)]^{1-\eta(v)}\left[\frac{l^n\left(v\right)}{\eta\left(v\right)}\right]^{\eta(v)}\left[\frac{l^r\left(v\right)}{1-\eta\left(v\right)}\right]^{1-\eta(v)}, \quad \text{(A.2)}$$

where $\omega^n(v) = w^o(v)$, $\omega^r(v) = \tau w^*(v)$ hold according to the main text and

$$\epsilon[\eta\left(v\right)] \equiv \frac{1-\eta(v)}{\int_{\eta(v)}^1 t\left(\tilde{\eta}\right)d\tilde{\eta}} \quad \text{(A.3)}$$

reflects the *average* productivity loss arising from the extra labour costs $t(\tilde{\eta})$, when producing a task abroad. Solving maximisation problem (A.2) gives marginal production costs $c\left(v\right) = w^o\left(v\right)/\left[\varphi\left(v\right)\tilde{z}\left(v\right)\right]$, where[2]

$$\tilde{z}(v) \equiv \left\{\frac{w^o(v)}{w^*(v)\tau}\epsilon[\eta\left(v\right)]\right\}^{1-\eta(v)}. \quad \text{(A.4)}$$

At stage one, the firm sets $\eta(v)$ to minimise its marginal cost $c(v)$. Thus, for the optimal $\eta(v)$-level the following first-order condition must hold: $\partial c(v)/\partial\eta(v) \overset{!}{=} 0$. In view of Eqs. (A.3)

[1] As in the main text, we define $l^r(v)$ such that foreign labour demand of offshoring firm v is given by $\tau l^r(v)$. While this definition of $l^r(v)$ might seem awkward at a first glance, it is useful for our purpose because it allows us to directly compare the production technology in Eq. (A.2) with the respective technology in Eq. (3.3).

[2] It is notable that $\tilde{z}(v)$ degenerates to $z(v)$, when considering a discrete offshoring technology, with

$$t(\tilde{\eta}) = \begin{cases} \infty & \forall\ \tilde{\eta} \in [0,\eta) \\ 1 & \forall\ \tilde{\eta} \in [\eta,1] \end{cases}.$$

and (A.4), this is equivalent to

$$\frac{\partial \ln \tilde{z}(v)}{\partial \eta(v)} = -\ln\left(\frac{w^o(v)}{w^*(v)\tau}\epsilon[\eta(v)]\right) + t[\eta(v)]\epsilon[\eta(v)] - 1 \stackrel{!}{=} 0. \tag{A.5}$$

Acknowledging Eq. (3.23) in the main text, we know that $w^o(v)/w^*(v)$ is the same for all producers, and hence Eq. (A.5) determines the same cost-minimising η for all firms. Since the second-order condition of the stage one cost-minimisation problem requires $\partial^2 \ln \tilde{z}(v)/\partial\eta(v)^2 < 0$, while $\partial^2 \ln \tilde{z}(v)/\partial\eta(v)\partial\tau > 0$ follows from inspection of Eq. (A.5), we can finally conclude that $d\eta/d\tau > 0$, and hence firms offshore a lower share of tasks if the costs of shipping foreign output back to the source country increase. This completes our formal discussion. ■

A.2 Derivation of $\gamma(\chi, \eta)$

We first show that the two averages $\bar{\pi}^o$ and $\bar{\pi}^d$ are proportional to $\pi^d(\varphi^d)$. An analogous result has already been shown in the main text for $\bar{\pi}$. It is an immediate implication of the Pareto distribution of productivities that the average operating profits of offshoring firms $\bar{\pi}^o$ are a multiple ζ of the marginal offshoring firm's operating profits $\pi^o(\varphi^o)$. Hence, we can write:

$$\bar{\pi}^o = \zeta\pi^o(\varphi^o) = \zeta\left[\frac{\pi^o(\varphi^o)}{\pi^d(\varphi^o)}\right]\left[\frac{\pi^d(\varphi^o)}{\pi^d(\varphi^d)}\right]\pi^d(\varphi^d) = \zeta\left(1+\chi^{-\frac{\xi}{k}}\right)\pi^d(\varphi^d), \tag{A.6}$$

where $\pi^o(\varphi^o)/\pi^d(\varphi^o) = 1 + \chi^{\xi/k}$ from Eq. (3.6) and the definition of χ reflects the firm-level productivity effect, while $\pi^d(\varphi^o)/\pi^d(\varphi^d) = (\varphi^d/\varphi^o)^{-\xi} = \chi^{-\xi/k}$ from Eq. (3.7) and the definition of χ captures the positive selection of offshoring firms. Using $\bar{\pi} = (1-\chi)\bar{\pi}^d + \chi\bar{\pi}^o$ as well as the solutions we have derived for $\bar{\pi}$ and $\bar{\pi}^o$ in terms of $\pi^d(\varphi^d)$, we get:

$$\bar{\pi}^d = \frac{\bar{\pi} - \chi\bar{\pi}^o}{1-\chi} = \zeta\frac{1-\chi^{\frac{k-\xi}{k}}}{1-\chi}\pi^d(\varphi^d). \tag{A.7}$$

Substituting for $\bar{\pi}$, $\bar{\pi}^o$, and $\bar{\pi}^d$ in the definition of γ, we then obtain $\gamma(\chi;\eta)$ as given in the main text. ■

A.3 Derivation of the Gini coefficient in Eq. (3.20)

For characterising the Gini coefficient in Eq. (3.20), we must distinguish between firms which offshore and those that produce only domestically. Cumulative profits of purely domestic firms with productivity $\bar{\varphi} \in [\varphi^d, \varphi^o)$ are given by $\Psi(\bar{\varphi}) \equiv N\int_{\varphi^d}^{\bar{\varphi}} \pi^d(\varphi)dG(\varphi)$. Considering $\pi^d(\varphi)/\pi^d(\varphi^d) = (\varphi/\varphi^d)^\xi$ from Eq. (3.7), we can solve for

$$\Psi(\bar{\varphi}) = M\pi^d(\varphi^d)\zeta\left[1 - \left(\frac{\bar{\varphi}}{\varphi^d}\right)^{\xi-k}\right]. \tag{A.8}$$

Economy-wide profit income is given by $\Psi = M(1+\chi)\zeta\pi^d(\varphi^d) - M\chi s$. Accounting for $s = \pi^d(\varphi^d)$ from Eq. (3.8), gives $\Psi = M\pi^d(\varphi^d)[\zeta + (\zeta-1)\chi]$. The share of cumulative profits realised by firms with a productivity level up to $\bar{\varphi} \in [\varphi^d, \varphi^o)$ is therefore given by

$$\frac{\Psi(\bar{\varphi})}{\Psi} = \frac{\zeta}{\zeta + (\zeta-1)\chi}\left[1 - \left(\frac{\bar{\varphi}}{\varphi^d}\right)^{\xi-k}\right]. \tag{A.9}$$

Denoting the fraction of firms with a productivity level $\varphi \leq \bar{\varphi}$ by $\mu \equiv 1(\bar{\varphi}/\varphi^d)^{-k}$, Eq. (A.9) can be rewritten as the first segment of the Lorenz curve for the distribution of profit income:

$$Q_M^1(\mu) = \frac{\zeta}{\zeta + (\zeta-1)\chi}\left[1 - (1-\mu)^{\frac{k-\xi}{k}}\right], \tag{A.10}$$

which is relevant for parameter domain $\mu \in [0, 1-\chi)$.

We now follow the same steps as above to calculate the second segment of the Lorenz curve for the distribution of profit income. We can first note that cumulative profits of all firms with productivities up to $\bar{\varphi} \in [\varphi^o, \infty)$ are given by $\Psi(\bar{\varphi}) = \Psi(\varphi^o) + N\int_{\varphi^o}^{\bar{\varphi}} \pi^o(\varphi)dG(\varphi) - N\int_{\varphi^o}^{\bar{\varphi}} sdG(\varphi)$. Accounting for $\pi^d(\varphi)/\pi^d(\varphi^d) = (\varphi/\varphi^d)^\xi$ from Eq. (3.7) and $\pi^o(\varphi)/\pi^d(\varphi) = 1 + \chi^{\xi/k}$, according to Eqs. (3.6) and (3.17), we can calculate

$$\Psi(\bar{\varphi}) = \Psi(\varphi^o) + M\pi^d(\varphi^d)\left\{\zeta\left(1+\chi^{\frac{\xi}{k}}\right)\left[\chi^{\frac{k-\xi}{k}} - \left(\frac{\bar{\varphi}}{\varphi^d}\right)^{\xi-k}\right] - \left[\chi - \left(\frac{\bar{\varphi}}{\varphi^d}\right)^{-k}\right]\right\}. \tag{A.11}$$

Dividing the latter by economy-wide profit income Ψ gives the share of profit income accruing to entrepreneurs with an ability up to $\bar{\varphi} \in [\varphi^o, \infty)$:

$$\frac{\Psi(\bar{\varphi})}{\Psi} = Q_M^1(1-\chi) + \frac{\zeta\left(1+\chi^{\frac{\xi}{k}}\right)\left[\chi^{\frac{k-\xi}{k}} - \left(\bar{\varphi}/\varphi^d\right)^{\xi-k}\right]}{\zeta + (\zeta-1)\chi} - \frac{\chi - \left(\bar{\varphi}/\varphi^d\right)^{-k}}{\zeta + (\zeta-1)\chi}. \tag{A.12}$$

Substituting μ from above, Eq. (A.12) can be reformulated to the second segment of the Lorenz curve, which is relevant for $\mu \in (1-\chi, 1]$

$$Q_M^2(\mu) = Q_M^1(1-\chi) + \frac{\zeta\left(1+\chi^{\frac{\xi}{k}}\right)\left[\chi^{\frac{k-\xi}{k}} - (1-\mu)^{\frac{k-\xi}{k}}\right]}{\zeta + (\zeta-1)\chi} - \frac{\mu - 1 + \chi}{\zeta + (\zeta-1)\chi}, \tag{A.13}$$

with $Q_M^2(1-\chi) = Q_M^1(1-\chi)$. Together Eqs. (A.10) and (A.13) form the Lorenz curve[3]

$$Q_M(\mu) \equiv \begin{cases} Q_M^1(\mu) & \text{if} \quad \mu \in [0, 1-\chi) \\ Q_M^2(\mu) & \text{if} \quad \mu \in [1-\chi, 1] \end{cases}. \tag{A.14}$$

The Gini coefficient for the distribution of profit income in Eq. (3.20) can then be computed according to $A_M(\chi) \equiv 1 - 2\int_0^1 Q_M(\mu)d\mu$. ■

A.4 Income inequality among self-employed agents

To characterise income inequality among *all* self-employed agents, we rely on the Lorenz curve for this income group, which now has three segments.[4] The first segment captures the share of income attributed to service providers. It is given by $Q_S^0(\mu) = \mu/\zeta$ and relevant for all $\mu \in [0, \chi/(1+\chi))$. The second segment of the Lorenz curve captures the income of service providers plus cumulative profits of purely domestic firms with a productivity level up to $\bar{\varphi} \in [\varphi^d, \varphi^o)$. Following the derivation steps in Appendix A.3, we can compute

$$\hat{\Psi}(\bar{\varphi}) = M\pi^d(\varphi^d)\left\{\chi + \zeta\left[1 - \left(\frac{\bar{\varphi}}{\varphi^d}\right)^{\xi-k}\right]\right\}. \tag{A.15}$$

Economy-wide profits plus service fees add up to total operating profits $\hat{\Psi} = M\pi^d(\varphi^d)(1+\chi)\left[k/(k-\xi)\right]$. Hence, the cumulative share of (profit) income realised by service providers and firms with a productivity level up to $\bar{\varphi} \in [\varphi^d, \varphi^o)$ is given by

$$\frac{\hat{\Psi}(\bar{\varphi})}{\hat{\Psi}} = \frac{1}{\zeta}\frac{\chi}{1+\chi} + \frac{1}{1+\chi}\left[1 - \left(\frac{\bar{\varphi}}{\varphi^d}\right)^{\xi-k}\right]. \tag{A.16}$$

We have to link Eq. (A.16) with the ratio of self-employed agents receiving the respective income share. Denoting the fraction of these agents by $\mu \equiv (1+\chi)^{-1}\left[1+\chi - (\bar{\varphi}/\varphi^d)^{-k}\right]$, Eq. (A.16) can be reformulated to the second segment of the Lorenz curve

$$Q_S^1(\mu) = \frac{1}{\zeta}\frac{\chi}{1+\chi} + \frac{1}{1+\chi}\left\{1 - \left[(1+\chi)(1-\mu)\right]^{\frac{k-\xi}{k}}\right\}, \tag{A.17}$$

[3] The Lorenz curve in Eq. (A.14) has the usual properties: $Q_M(0) = 0$, $Q_M(1) = 1$ and $Q_M'(\mu) > 0 \;\; \forall \;\; \mu \in (0,1)$.

[4] In this subsection, we consider the basic model variant without rent sharing. The respective results for the model variant with rent sharing are obtained when replacing ξ by $\bar{\xi}$.

which is relevant for parameter domain $\mu \in [\chi/(1+\chi), 1/(1+\chi))$.

In a final step, we compute the cumulative income of all service providers and entrepreneurs with an ability up to $\bar{\varphi} \in [\varphi^o, \infty)$ as a share of the total income of self-employed agents, $\hat{\Psi}$. Substituting μ from above, this gives the third segment of the Lorenz curve

$$Q_S^2(\mu) = Q_S^1\left(\frac{1}{1+\chi}\right) + \frac{1}{1+\chi}\left\{\left(1+\chi^{\xi/k}\right)\left[\chi^{\frac{k-\xi}{k}} - [(1+\chi)(1-\mu)]^{\frac{k-\xi}{k}}\right]\right.$$
$$\left. -\frac{1}{\zeta}[\chi - (1+\chi)(1-\mu)]\right\}. \qquad \text{(A.18)}$$

Putting the three segments together, we obtain the new Lorenz curve

$$Q_S(\mu) \equiv \begin{cases} Q_S^0(\mu) & \text{if} \quad \mu \in \left[0, \frac{\chi}{1+\chi}\right) \\ Q_S^1(\mu) & \text{if} \quad \mu \in \left[\frac{\chi}{1+\chi}, \frac{1}{1+\chi}\right) \\ Q_S^2(\mu) & \text{if} \quad \mu \in \left[\frac{1}{1+\chi}, 1\right] \end{cases}. \qquad \text{(A.19)}$$

The Gini coefficient for the distribution of income among self-employed agents can then be calculated according to $A_S(\chi) \equiv 1 - 2\int_0^1 Q_S(\mu)d\mu$. Substituting Eq. (A.19), we can compute the respective expression in Footnote 19. ■

A.5 Proof of Proposition 3.2.3

Substitution of Eqs. (3.14) and (3.16) for M and φ^d, respectively, in Eq. (3.22) and using the resulting expression in $\Phi(\chi) = I(\chi)/I(0)$, we obtain after tedious but straightforward computations: $\Phi(\chi) = T_1(\chi) \times T_2(\chi) \times T_3(\chi)$, with

$$T_1(\chi) \equiv \frac{[1+\gamma(\sigma-1)]\,[1+\zeta(\sigma-1)]}{\sigma\,[1+\gamma\zeta(\sigma-1)]}, \qquad T_2(\chi) \equiv \left\{\frac{(1+\chi)\,[1+\gamma\zeta(\sigma-1)]}{1+\zeta(\sigma-1)}\right\}^{\frac{\sigma-1-\varepsilon k}{k(\sigma-1)}}, \quad \text{(A.20)}$$

and $T_3(\chi) \equiv (1+\chi)^{1/(\sigma-1)}$. Differentiation of $\Phi(\chi)$ establishes

$$\Phi'(\chi) = \frac{\Phi(\chi)}{1+\gamma(\sigma-1)}\left\{-\left[\frac{\hat{\kappa}(\chi;\varepsilon)+\xi(\sigma-1)}{\gamma k(\sigma-1)+k-\xi}\right]\frac{\partial\gamma}{\partial\chi} + \left[\frac{(1-\varepsilon)\,k+\sigma-1}{k(\sigma-1)}\right]\frac{1+\gamma(\sigma-1)}{1+\chi}\right\}, \qquad \text{(A.21)}$$

with $\hat{\kappa}(\chi;\varepsilon) \equiv [k\varepsilon-\sigma+1]\,[1+\gamma(\chi;\eta)\,(\sigma-1)]$. In view of $\partial\gamma/\partial\chi < 0$, it is immediate that $\hat{\kappa}(\chi;\varepsilon)+\xi(\sigma-1) \geq 0$ is sufficient for $\Phi'(\chi) > 0$. $\Phi'(\chi) > 0$ is therefore guaranteed if $\varepsilon \geq (\sigma-1)/k$, because in this case we have $\hat{\kappa}(\chi;\varepsilon) \geq 0$. Things are less obvious for parameter domain $\varepsilon <$

$(\sigma-1)/k$, because in this case we have $\hat{\kappa}(\chi;\varepsilon)<0$. However, noting that for parameter domain $\varepsilon<(\sigma-1)/k$ we have $\partial\hat{\kappa}(\chi;\varepsilon)/\partial\chi>0$, it follows that in this case $\hat{\kappa}(\chi;\varepsilon)+\xi(\sigma-1)>0$ must hold for all possible χ if $\hat{\kappa}(0;\varepsilon)+\xi(\sigma-1)\geq 0$ or, equivalently, if $\varepsilon\geq(\sigma-\xi)(\sigma-1)/(\sigma k)\equiv\bar{\varepsilon}$. We can thus safely conclude that $\Phi'(\chi)>0$ is guaranteed if $\varepsilon\geq\bar{\varepsilon}$. Put differently, if $\varepsilon\geq\bar{\varepsilon}$ source country welfare is monotonically increasing in the share of offshoring firms, and hence welfare in the source country is unambiguously higher with offshoring than in autarky.

We now consider the parameter domain $\varepsilon<\bar{\varepsilon}$. In this case, $\Phi'(0)<0$ follows from $\hat{\kappa}(0;\varepsilon)+\xi(\sigma-1)<0$ and the fact that $\lim_{\chi\to 0}\partial\gamma/\partial\chi=-\infty$, and hence offshoring lowers source country welfare relative to autarky if χ is small. Furthermore, evaluating the derivative in Eq. (A.21) at $\chi=1$, we obtain

$$\Phi'(1)=\frac{\Phi(1)}{2k}\,\frac{[b(\eta)+k(k-\xi)(1-\eta)]\,[1+\eta(\sigma-1)]+\xi(\sigma-1)(k-\xi)(1-\eta)}{[1+\eta(\sigma-1)]\,[\eta k(\sigma-1)+k-\xi]}, \tag{A.22}$$

with $b(\eta)\equiv[(1-\varepsilon)k+\sigma-1]\,[(2\eta-1)k+\xi(1-\eta)+(k-\xi)/(\sigma-1)]$. It is immediate that $\eta>0.5$ is sufficient for $b(\eta)>0$ and in this case we have $\Phi'(1)>0$. Hence, if $\eta>0.5$, offshoring exerts a non-monotonic effect on source country welfare. Noting that $T_1(\chi)>1$ and $T_3(\chi)>1$ hold for any $\chi>0$, whereas $\eta>0.5$ is sufficient for $T_2(1)>1$ if $\varepsilon<(\sigma-1)/k$, we can safely conclude that $\Phi(1)>1$, and hence the source country benefits from offshoring if χ is large.

Taking stock, our previous analysis has identified a critical level of external increasing returns to scale in the production of Y: $\bar{\varepsilon}$. If external increasing returns to scale are larger than the critical level, offshoring exerts a positive monotonic effect on source country welfare. In contrast, if the external increasing returns to scale are smaller than the critical level, offshoring exerts a non-monotonic effect on source country welfare. In this case, the source country is worse off with offshoring than under autarky if χ is small, while it benefits from offshoring if χ is large. Using the parameter estimates from Section 3.5, we can determine 0.61 as an empirically plausible value for $\bar{\varepsilon}$. Empirical estimates for parameter ε are reported by Ardelean (2011). Relying on UN COMTRADE data, Ardelean identifies an average value of $\varepsilon=0.56$ across all industries in her data-set. This lends support to the non-monotonic welfare effect of offshoring in Proposition 3.2.3.[5] This completes the proof. ■

[5] Ardelean (2011) does not distinguish between final and intermediate goods, and hence her ε estimates capture

A.6 The social planner problem for $\varepsilon = 1$ under autarky

In autarky, the social planner sets φ^d and the quantity $q(v) > 0$ of all varieties v to maximize output Y, subject to the binding resource constraint. We first consider the problem of setting optimal quantities $q(v)$ for a given φ^d. Holding φ^d constant under autarky is tantamount to fixing the amount of resources used as variable production input: $L = NG(\varphi^d)$. The social planner's problem in this case is therefore to maximize $Y = \left[\int_{v\in V} q(v)^\rho dv\right]^{1/\rho}$, subject to $\int_{v\in V} [q(v)/\varphi(v)]dv = NG(\varphi^d)$. The first-order conditions for this maximization problem establish for any two varieties v_1, v_2 the following output ratio: $q(v_1)/q(v_2) = [\varphi(v_1)/\varphi(v_2)]^\sigma$. This implies that output increases with productivity and hence, we can refer to varieties by means of the underlying productivity parameter. The marginal variety is the one with the lowest output and produced with productivity φ^d. We can define $a \equiv q(\varphi^d)/(\varphi^d)^\sigma$. An optimal allocation of resources then requires that the output level of any firm with productivity $\varphi \geq \varphi^d$ is set to $q(\varphi) = a\varphi^\sigma$, with $a > 0$.

With these insights at hand, we can reformulate the social planner's problem as

$$\max_{\varphi^d, a} \; Y = \left[N\int_{\varphi^d}^{\infty} q(\varphi)^\rho \, dG(\varphi)\right]^{\frac{1}{\rho}} \quad \text{s.t.} \quad \int_{\varphi^d}^{\infty}\left[\frac{q(\varphi)}{\varphi}\right] dG(\varphi) = G(\varphi^d), \quad q(\varphi) = a\varphi^\sigma. \quad \text{(A.23)}$$

Applying $q(\varphi) = a\varphi^\sigma$, we can rewrite the resource constraint as follows: $\zeta a(\varphi^d)^{\sigma-1-k} = 1 - (\varphi^d)^{-k}$. Furthermore, economy-wide output can be written as $Y = \left[N\zeta a^\rho(\varphi^d)^{\sigma-1-k}\right]^{1/\rho}$. Solving the resource constraint for a and substituting the resulting expression into Y, we can simplify the social planner's problem to

$$\max_{\varphi^d} \; N^{\frac{\sigma}{\sigma-1}}\zeta^{\frac{1}{\sigma-1}}\left[1-\left(\varphi^d\right)^{-k}\right]\left(\varphi^d\right)^{\frac{\sigma-1-k}{\sigma-1}}. \quad \text{(A.24)}$$

The first-order condition to this maximization problem establishes $\varphi^d = \left[1+\zeta(\sigma-1)\right]^{1/k}$ and this coincides with the outcome of decentralized firm entry in Eq. (3.16), when considering $\chi = 0$. Hence, for $\varepsilon = 1$ the market equilibrium under autarky is allocationally efficient. ■

external increasing returns to scale due to a love of variety of consumers in a Krugman-type model as well as external increasing returns to scale in final goods production due to labor division in an Ethier framework. Furthermore, it is notable that variation in the ε estimates is large, ranging from a low level of 0.19 in the 'Headgear and Parts Thereof' industry to a relatively high level of 0.88 in the 'Soap etc.; Waxes, Polish, etc; Candles; Dental Preps' industry.

A.7 Derivation details for the model variant with $\theta > 0$

In this subsection, we show in detail how the equations in Section 3.1 must be modified, when allowing for rent sharing between workers and firms. The first equation that has to be modified is Eq. (3.4). With rent sharing wages are firm-specific, and hence we can rewrite marginal production costs as follows:

$$c(v) = \frac{\omega^n(v)}{\varphi(v)z(v)} \qquad \text{with} \qquad z(v) \equiv \left[\frac{\omega^n(v)}{\omega^r(v)}\right]^{1-\eta}, \tag{A.25}$$

where $\omega^n(v)$ is the domestic wage rate paid by firm v to workers conducting non-routine tasks. Thereby, we have $\omega^n(v) = w^d(v)$ if the firm produces all tasks at home, while we have $\omega^n(v) = w^o(v)$ if routine tasks are produced offshore. As in Section 3.1, we have $z(v) = 1$ and thus $c^d(v) = w^d(v)/\varphi$ if the firm produces purely domestically. For an offshoring firm, we obtain $z(v) = z^o(v)$ and, instead of Eq. (3.4),

$$c^o(v) = \frac{w^o(v)}{\varphi(v)z^o(v)}, \qquad \text{where} \qquad z^o(v) \equiv \left[\frac{w^o(v)}{\tau w^*(v)}\right]^{1-\eta} = \left[\frac{(1-U)\,\bar{w}}{(1-U^*)\bar{w}^*}\right]^{(1-\eta)(1-\theta)} \tau^{\eta-1}. \tag{A.26}$$

Thereby, we have made use of the fair-wage constraint in Eq. (3.23) in order to substitute for $w^o(v)/w^*(v)$. Combining Eqs. (3.6) and (3.23), we can furthermore compute

$$\frac{\pi^o(v)}{\pi^d(v)} = \kappa^{(\sigma-1)} \qquad \text{and} \qquad \frac{w^o(v)}{w^d(v)} = \kappa^{\theta(\sigma-1)}, \tag{A.27}$$

where

$$\kappa \equiv \frac{c^d(v)}{c^o(v)} = \left\{\left[\frac{(1-U)\,\bar{w}}{(1-U^*)\bar{w}^*}\right]^{(1-\eta)(1-\theta)} \left(\frac{1}{\tau}\right)^{1-\eta}\right\}^{\frac{\bar{\xi}}{\sigma-1}}. \tag{A.28}$$

Using Eqs. (3.7) and (A.27) in indifference condition (3.9), and accounting for $\pi^d(\varphi^d) = s$ from Eq. (3.8′), we can easily verify that the link between χ and κ continues to be given by Eq. (3.17), with $\bar{\xi}$ assuming the role of ξ if $\theta > 0$. Labour income per capita in the source and host country are given by

$$(1-U)\bar{w} = \frac{\gamma\rho Y}{L} \qquad \text{and} \qquad (1-U^*)\bar{w}^* = \frac{(1-\gamma)\,\rho Y}{N^*}, \tag{A.29}$$

respectively. Substituting Eqs. (A.29) and (3.15) into Eq. (A.28) allows us to compute

$$\kappa = \left\{\left[\frac{\gamma k(\sigma-1) + k - \bar{\xi}}{(1-\gamma)\,k(\sigma-1)}\left(\frac{N^*}{N}\right)\right]^{(1-\eta)(1-\theta)} \left(\frac{1}{\tau}\right)^{1-\eta}\right\}^{\frac{\bar{\xi}}{\sigma-1}}. \tag{A.30}$$

And combining Eqs. (3.17) and (A.30) we can conclude that the relationship between κ and χ in the model variant with $\theta > 0$ is characterised by the implicit function

$$F(\chi,\tau) \equiv \left\{ \left[\frac{\gamma k(\sigma-1)+k-\bar{\xi}}{(1-\gamma)\,k(\sigma-1)} \left(\frac{N^*}{N}\right) \right]^{(1-\eta)(1-\theta)} \left(\frac{1}{\tau}\right)^{1-\eta} \right\}^{\frac{\bar{\xi}}{\sigma-1}} - \left(1+\chi^{\frac{\bar{\xi}}{k}}\right)^{\frac{1}{\sigma-1}} = 0.$$

This completes our discussion on how rent sharing affects the equations reported in Section 3.1. ■

A.8 Derivation of Eq. (3.24)

Adding up domestic employment over all purely domestic and offshoring firms in the source country gives $(1-U)\,L = N\left[\int_{\varphi^d}^{\varphi^o} l^d(\varphi)\,dG(\varphi) + \int_{\varphi^o}^{\infty} l^o(\varphi)\,dG(\varphi)\right]$. Using $l^d(\varphi)\,/l^d(\varphi^d) = (\varphi/\varphi^d)^{(1-\theta)\bar{\xi}}$ and $l^o(\varphi)\,/l^d(\varphi^d) = \eta\kappa^{(\sigma-1)(1-\theta)}(\varphi/\varphi^d)^{(1-\theta)\bar{\xi}}$, according to Eqs. (3.2), (3.7), the equivalent of Eq. (3.19) for the scenario with $\theta > 0$, and Eq. (3.23), and accounting for the definition of $\beta(\chi;\eta)$ in Eq. (3.25), we can calculate

$$(1-U)L = Ml^d(\varphi^d)\beta(\chi;\eta)\frac{\bar{\zeta}}{1+\theta(\bar{\zeta}-1)}. \tag{A.31}$$

Furthermore, combining Eqs. (3.8′), (3.10), (3.11) and noting that constant markup pricing implies $(\sigma-1)\pi(\varphi^d) = l^d(\varphi^d)w(\varphi^d)$, we can express the total wage bill in the source country as follows:

$$(1-U)L\bar{w} = Ml^d(\varphi^d)w(\varphi^d)\alpha(\chi;\eta)\bar{\zeta}. \tag{A.32}$$

Together Eq. (A.31) and Eq. (A.32) determine the wage ratio $w(\varphi^d)/\bar{w} = \Delta(\chi;\eta)/[1+\theta(\bar{\zeta}-1)]$, where $\Delta(\chi;\eta) = \beta(\chi;\eta)/\alpha(\chi;\eta)$ has been considered. Applying the fair-wage constraint (3.23) for the marginal firm and considering indifference condition (3.8), we can compute $U = 1 - w(\varphi^d)/\bar{w}$. Substituting for $w(\varphi^d)/\bar{w}$, then gives Eq. (3.24). ■

A.9 Proof of Proposition 3.3.1

Let us first consider the impact of offshoring on U. From Eq. (3.25), we can conclude that, for all $\chi \in (0,1]$, $\Delta(\chi;\eta) >,=,< 1$ is equivalent to $\Omega(\vartheta) \equiv (\eta\vartheta^{1-\theta}-1)\,(\vartheta-1)^{\theta} - (\eta\vartheta-1) >,=,< 0$, with $\vartheta \equiv 1+\chi^{\bar{\xi}/k} \in (1,2]$. Differentiating $\Omega(\vartheta)$ gives $\Omega'(\vartheta) = -\eta[1-(1-\theta)\vartheta^{-\theta}(\vartheta-1)^{\theta}] +$

$\theta\,(\vartheta-1)^{\theta-1}\,(\eta\vartheta^{1-\theta}-1)$ and $\Omega''(\vartheta)=\theta(1-\theta)(\vartheta-1)^{\theta-2}[1-\eta/\vartheta^{1+\theta}]$. Accounting for $\Omega''(\vartheta)>0$ and $\Omega'(2)=-\eta(1-2^{-\theta})-\theta(1-\eta 2^{-\theta})<0$, it follows that $\Omega'(\vartheta)<0$ must hold for all $\vartheta\in(1,2)$. Noting finally that $\Omega(1)=1-\eta>0$ and $\Omega(2)=-2\eta[1-(1/2)^{\theta}]<0$, we can define a unique $\hat{\chi}\in(0,1)$, such that offshoring lowers U if $\chi<\hat{\chi}$, while it raises U if $\chi>\hat{\chi}$.

From inspection of Eq. (3.26) we can note that $\Lambda>1$ requires $\Delta<1$ and thus a positive effect of offshoring on U. This implies that $\Lambda(\chi;\eta)>1$ can only materialise if $\chi>\hat{\chi}$. Furthermore, it is worth noting that partially differentiating $\Delta(\chi;\eta)$ with respect to η gives

$$\frac{\partial\Delta\left(\chi;\eta\right)}{\partial\eta}=-\frac{\chi^{\frac{k-\bar{\xi}}{k}}\left(1+\chi^{\frac{\bar{\xi}}{k}}\right)\left[\left(1-\chi^{\frac{k-(1-\theta)\bar{\xi}}{k}}\right)-\left(1-\chi^{\frac{k-\bar{\xi}}{k}}\right)\chi^{\frac{\theta\bar{\xi}}{k}}\left(1+\chi^{\frac{\bar{\xi}}{k}}\right)^{-\theta}\right]}{\alpha(\chi;\eta)^{2}}<0. \quad \text{(A.33)}$$

Additionally accounting for $\partial\gamma\left(\chi;\eta\right)/\partial\eta>0$, it follows from Eq. (3.26) that $\partial\Lambda\left(\chi;\eta\right)/\partial\eta>0$. Considering $\Lambda(1;0)=0$ and $\Lambda(1;1)>1$, this implies that $\Lambda(1;\eta)=1$ has a unique solution in $\eta\in(0,1)$, which is given by $\hat{\eta}$ in Proposition 3.3.1. We can thus safely conclude that u/u^{a} is non-monotonic in χ if $\eta>\hat{\eta}$, with $u/u^{a}<1$ if χ sufficiently small and $u/u^{a}>1$ if χ close to one.

We finally show that $\hat{\eta}<0.5$ if $k\geq 2$. For this purpose, we can note that $\hat{\eta}>,=,<0.5$ is equivalent to $\Gamma(\theta,k,\bar{\xi})\equiv 2^{\theta}\theta\bar{\xi}-(2^{\theta}-1)(k\sigma-\bar{\xi})>,=,<0$. To determine the sign of $\Gamma(\theta,k,\bar{\xi})$, let us first consider a parameter domain with $k\geq\bar{\xi}\geq 2$. In this case, we have $\Gamma(\theta,k,\bar{\xi})\leq\Gamma(\theta,\bar{\xi},\bar{\xi})=g(\theta)\bar{\xi}$, with $g(\theta)\equiv(\sigma-1)-2^{\theta}(\sigma-1-\theta)$. Since $\bar{\xi}\geq 2$ implies $\sigma-1\geq 2$ and $g(\theta)$ decreases in $\sigma-1$, we can further conclude that $g(\theta)\leq 2-2^{\theta}(2-\theta)\equiv\underline{g}(\theta)$. Differentiation of $\underline{g}(\theta)$ gives $\underline{g}'(\theta)=2^{\theta}\left[1-\ln 2(2-\theta)\right]$ and $\underline{g}''(\theta)=2^{\theta}\ln 2\left[2-\ln 2(2-\theta)\right]>0$. From inspection of these derivatives, it follows that $\underline{g}(\theta)$ has a unique extremum, which is a minimum. Noting further that $\underline{g}(0)=\underline{g}(1)=0$, it is clear that $\underline{g}(\theta)<0$ must hold for all $\theta\in(0,1)$. This proves that $\Gamma(\theta,k,\bar{\xi})<0$ and thus $\hat{\eta}<0.5$ if $k\geq\bar{\xi}\geq 2$. Let us now consider a parameter domain with $k\geq 2>\bar{\xi}$. In this case, we have $\Gamma(\theta,k,\bar{\xi})\leq\Gamma(\theta,2,\bar{\xi})=\hat{g}(\theta)\bar{\xi}$, with $\hat{g}(\theta)\equiv 2^{\theta}\theta-(2^{\theta}-1)[(\sigma+1)/(\sigma-1)+2\sigma\theta]$. Noting that $\sigma+1+2\sigma\theta(\sigma-1)=(\sigma-1)[1+2/\bar{\xi}+2\theta(\sigma-1)]$, it follows from $\bar{\xi}<2$ that $\sigma+1+2\sigma\theta(\sigma-1)>2(\sigma-1)[1+\theta(\sigma-1)]$ and thus $\hat{g}(\theta)<2[1+\theta(\sigma-1)]-2^{\theta}\{2[1+\theta(\sigma-1)]-\theta\}<\underline{g}(\theta)$. From above, we know that $\underline{g}(\theta)<0$ holds for all $\theta\in(0,1)$. This confirms that $\Gamma(\theta,k,\bar{\xi})<0$ and thus $\hat{\eta}<0.5$ if $k\geq 2>\bar{\xi}$, which completes the proof. ∎

A.10 Derivation of the Gini coefficient in Eq. (3.27)

To characterise the Gini coefficient for the distribution of wage income we must distinguish workers employed in purely domestic firms from those employed in offshoring firms. Cumulative labour income of workers employed in purely domestic firms with a productivity level up to $\bar{\varphi} \in [\varphi^d, \varphi^o)$ is given by $W(\bar{\varphi}) \equiv N \int_{\varphi^d}^{\bar{\varphi}} w^d(\varphi) l^d(\varphi) dG(\varphi)$. Since constant markup pricing implies that a firm's wage bill is proportional to its revenues, we can make use of $w^d(\varphi)l^d(\varphi) = (\sigma - 1)\pi^d(\varphi)$. Considering $\pi^d(\varphi)/\pi^d(\varphi^d) = (\varphi/\varphi^d)^{\bar{\xi}}$ from Eq. (3.7), then gives

$$W(\bar{\varphi}) = (\sigma - 1)M\pi^d(\varphi^d)\bar{\zeta}\left[1 - \left(\frac{\bar{\varphi}}{\varphi^d}\right)^{\bar{\xi}-k}\right]. \tag{A.34}$$

Total economy-wide labour income equals $W = \rho\gamma Y$. Using Eq. (3.10) and the definition of γ, we obtain $W = (\sigma - 1)M\pi^d(\varphi^d)\bar{\zeta}\alpha(\chi;\eta)$. Hence, the share of wage income accruing to workers, who are employed in firms with a productivity level up to $\bar{\varphi} \in [\varphi^d, \varphi^o)$, can be expressed as

$$\frac{W(\bar{\varphi})}{W} = \frac{1}{\alpha(\chi;\eta)}\left[1 - \left(\frac{\bar{\varphi}}{\varphi^d}\right)^{\bar{\xi}-k}\right]. \tag{A.35}$$

To calculate the Lorenz curve for the distribution of labour income, we must link the income ratio in Eq. (A.35) to the respective employment ratio. For this purpose, we first note that total employment in all firms with a productivity level up to $\bar{\varphi} \in [\varphi^d, \varphi^o)$ is given by $L(\bar{\varphi}) \equiv N \int_{\varphi^d}^{\bar{\varphi}} l^d(\varphi) dG(\varphi)$. Substituting $l^d(\varphi)/l^d(\varphi^d) = (\varphi/\varphi^d)^{(1-\theta)\bar{\xi}}$, we can calculate

$$L(\bar{\varphi}) = Ml^d(\varphi^d)\frac{\bar{\zeta}}{1 + \theta(\bar{\zeta} - 1)}\left[1 - \left(\frac{\bar{\varphi}}{\varphi^d}\right)^{(1-\theta)\bar{\xi}-k}\right]. \tag{A.36}$$

In a similar vein, we can show that economy-wide employment of production workers in the source country equals $(1-U)L = Ml^d(\varphi^d)\beta(\chi;\eta)\bar{\zeta}/[1+\theta(\bar{\zeta}-1)]$. Hence, the share of production workers that are employed in firms with a productivity level up to $\bar{\varphi} \in [\varphi^d, \varphi^o)$ is given by $\lambda = \beta(\chi;\eta)^{-1}[1 - (\bar{\varphi}/\varphi^d)^{(1-\theta)\bar{\xi}-k}]$. Combining the latter with Eq. (A.35), we obtain the first segment of the Lorenz curve for the distribution of labour income

$$Q_L^1(\lambda) = \frac{1 - [1 - \beta(\chi;\eta)\lambda]^{\frac{k-\bar{\xi}}{k-(1-\theta)\bar{\xi}}}}{\alpha(\chi;\eta)}, \tag{A.37}$$

which is relevant if $\lambda \in [0, b_L)$, with $b_L \equiv \beta(\chi;\eta)^{-1}(1 - \chi^{1-(1-\theta)\bar{\xi}/k})$.

We now follow the same steps as above to calculate the second segment of the Lorenz curve. We first compute the total domestic wage bill of firms with a productivity level up to $\bar{\varphi} \in [\varphi^o, \infty)$. This gives $W(\bar{\varphi}) \equiv W(\varphi^o) + N \int_{\varphi^o}^{\bar{\varphi}} w^o(\varphi) l^o(\varphi) dG(\varphi)$. Accounting for $w^o(\varphi)l^o(\varphi) = \eta(\sigma-1)\pi^o(\varphi)$ and considering $\pi^d(\varphi)/\pi^d(\varphi^d) = (\varphi/\varphi^d)^{\bar{\xi}}$ from Eq. (3.7) as well as $\pi^o(\varphi)/\pi^d(\varphi) = 1 + \chi^{\bar{\xi}/k}$, according to Eqs. (3.6) and (3.17), we can calculate

$$W(\bar{\varphi}) = W(\varphi^o) + M\pi^d(\varphi^d)\bar{\zeta}(\sigma-1)\eta\left(1+\chi^{\frac{\bar{\xi}}{k}}\right)\left[\chi^{\frac{k-\bar{\xi}}{k}} - \left(\frac{\bar{\varphi}}{\varphi^d}\right)^{\bar{\xi}-k}\right]. \tag{A.38}$$

Dividing Eq. (A.38) by economy-wide labour income W, yields

$$\frac{W(\bar{\varphi})}{W} = 1 - \frac{\eta\left(1+\chi^{\frac{\bar{\xi}}{k}}\right)}{\alpha(\chi;\eta)}\left(\frac{\bar{\varphi}}{\varphi^d}\right)^{\bar{\xi}-k}. \tag{A.39}$$

The mass of domestic workers employed by firms with a productivity level up to $\bar{\varphi} \in [\varphi^o, \infty)$ is given by $L(\bar{\varphi}) = L(\varphi^o) + N \int_{\varphi^o}^{\bar{\varphi}} l^o(\varphi) dG(\varphi)$. Accounting for $l^o(\varphi)/l^d(\varphi) = \eta\kappa^{(\sigma-1)(1-\theta)} = \eta(1+\chi^{\bar{\xi}/k})^{1-\theta}$ and $l^d(\varphi)/l^d(\varphi^d) = (\varphi/\varphi^d)^{(1-\theta)\bar{\xi}}$, we can further write

$$L(\bar{\varphi}) = L(\varphi^o) + Ml^d(\varphi^d)\frac{\bar{\zeta}\eta\left(1+\chi^{\frac{\bar{\xi}}{k}}\right)^{1-\theta}}{1+\theta(\bar{\zeta}-1)}\left[\chi^{\frac{k-(1-\theta)\bar{\xi}}{k}} - \left(\frac{\bar{\varphi}}{\varphi^d}\right)^{(1-\theta)\bar{\xi}-k}\right]. \tag{A.40}$$

Dividing $L(\bar{\varphi})$ by economy-wide employment $(1-U)L$, then gives $\lambda = 1 - \eta\beta(\chi;\eta)^{-1}(1 + \chi^{\bar{\xi}/k})^{1-\theta}(\bar{\varphi}/\varphi^d)^{(1-\theta)\bar{\xi}-k}$. Solving the latter for $\bar{\varphi}/\varphi^d$ and substituting the resulting expression into Eq. (A.39), we obtain the second segment of the Lorenz curve

$$Q_L^2(\lambda) = 1 - \frac{\eta\left(1+\chi^{\frac{\bar{\xi}}{k}}\right)}{\alpha(\chi;\eta)}\left[\frac{(1-\lambda)\beta(\chi;\eta)}{\eta\left(1+\chi^{\frac{\bar{\xi}}{k}}\right)^{1-\theta}}\right]^{\frac{k-\bar{\xi}}{k-(1-\theta)\bar{\xi}}}, \tag{A.41}$$

which is relevant if $\bar{\varphi} \in [\varphi^o, \infty)$. Together Eqs. (A.37) and (A.41) form the Lorenz curve[6]

$$Q_L(\lambda) \equiv \begin{cases} Q_L^1(\lambda) & \text{if} \quad \lambda \in [0, b_L) \\ Q_L^2(\lambda) & \text{if} \quad \lambda \in [b_L, 1] \end{cases}. \tag{A.42}$$

[6]The Lorenz curve in Eq. (A.42) has the usual properties: $Q_L(0) = 0$, $Q_L(1) = 1$ and $Q_L'(\lambda) > 0 \;\forall\; \lambda \in (0,1)$.

The Gini coefficient for the distribution of labour income in Eq. (3.27) can then be computed according to $A_L(\chi) \equiv 1 - 2\int_0^1 Q_L(\lambda)d\lambda$. Thereby we can note that

$$\int_0^{b_L} Q_L^1(\lambda)d\lambda = \frac{1}{\alpha(\chi;\eta)}\left[\lambda + \frac{[1-\beta(\chi;\eta)\lambda]^{\frac{2(k-\bar{\xi})+\theta\bar{\xi}}{k-(1-\theta)\bar{\xi}}}}{\beta(\chi;\eta)}\frac{1+\theta(\bar{\zeta}-1)}{2+\theta(\bar{\zeta}-1)}\right]_0^{b_L}$$

$$= \frac{b_L}{\alpha(\chi;\eta)} + \frac{1+\theta(\bar{\zeta}-1)}{2+\theta(\bar{\zeta}-1)}\left[\frac{\chi^{\frac{2(k-\bar{\xi})+\theta\bar{\xi}}{k}}}{\alpha(\chi;\eta)\beta(\chi;\eta)} - \frac{1}{\alpha(\chi;\eta)\beta(\chi;\eta)}\right], \tag{A.43}$$

while

$$\int_{b_L}^1 Q_L^2(\lambda)d\lambda = \left[\lambda + \frac{\eta(1+\chi^{\bar{\xi}/k})}{\alpha(\chi;\eta)}\left[\frac{\beta(\chi;\eta)}{\eta(1+\chi^{\bar{\xi}/k})^{1-\theta}}\right]^{\frac{k-\bar{\xi}}{k-(1-\theta)\bar{\xi}}}\frac{1+\theta(\bar{\zeta}-1)}{2+\theta(\bar{\zeta}-1)}(1-\lambda)^{\frac{2(k-\bar{\xi})+\theta\bar{\xi}}{k-(1-\theta)\bar{\xi}}}\right]_{b_L}^1$$

$$= 1 - b_L - \frac{\eta^2(1+\chi^{\bar{\xi}/k})^{2-\theta}}{\alpha(\chi;\eta)\beta(\chi;\eta)}\frac{1+\theta(\bar{\zeta}-1)}{2+\theta(\bar{\zeta}-1)}\chi^{\frac{2(k-\bar{\xi})+\theta\bar{\xi}}{k}}. \tag{A.44}$$

Substituting Eqs. (A.43) and (A.44) into

$$A_L = 1 - 2\int_0^{b_L} Q_L^1(\lambda)d\lambda - 2\int_{b_L}^1 Q_L^2(\lambda)d\lambda, \tag{A.45}$$

we obtain

$$A_L = -1 + 2b_L\frac{\alpha(\chi;\eta)-1}{\alpha(\chi;\eta)} + \frac{2}{\alpha(\chi;\eta)\beta(\chi;\eta)}\frac{1+\theta(\bar{\zeta}-1)}{2+\theta(\bar{\zeta}-1)} + 2Z(\chi;\eta), \tag{A.46}$$

with

$$Z(\chi;\eta) \equiv \left[\frac{\eta^2\left(1+\chi^{\bar{\xi}/k}\right)^{2-\theta}}{\alpha(\chi;\eta)\beta(\chi;\eta)} - \frac{1}{\alpha(\chi;\eta)\beta(\chi;\eta)}\right]\frac{1+\theta(\bar{\zeta}-1)}{2+\theta(\bar{\zeta}-1)}\chi^{\frac{2(k-\bar{\xi})+\theta\bar{\xi}}{k}}. \tag{A.47}$$

Using the definition of b_L, we can rewrite A_L in the following way

$$A_L = \frac{\theta(\bar{\zeta}-1)}{2+\theta(\bar{\zeta}-1)} + 2\frac{1+\theta(\bar{\zeta}-1)}{2+\theta(\bar{\zeta}-1)}\frac{[1-\alpha(\chi;\eta)\beta(\chi;\eta)]}{\alpha(\chi;\eta)\beta(\chi;\eta)} + \frac{2\left(1-\chi^{\frac{k-(1-\theta)\bar{\xi}}{k}}\right)[\alpha(\chi;\eta)-1]}{\alpha(\chi;\eta)\beta(\chi;\eta)} + 2Z(\chi;\eta). \tag{A.48}$$

Accounting for Eq. (3.25), we can show that

$$1-\alpha(\chi;\eta)\beta(\chi;\eta) = -[\alpha(\chi;\eta)-1] - [\beta(\chi;\eta)-1] - [\alpha(\chi;\eta)-1][\beta(\chi;\eta)-1]$$

$$= -[\alpha(\chi;\eta)-1]\left(1-\chi^{\frac{k-(1-\theta)\bar{\xi}}{k}}\right) - [\beta(\chi;\eta)-1]\left(1-\chi^{\frac{k-\bar{\xi}}{k}}\right) - \left[\eta^2\left(1+\chi^{\bar{\xi}/k}\right)^{2-\theta} - 1\right]\chi^{\frac{2(k-\bar{\xi})+\theta\bar{\xi}}{k}}. \tag{A.49}$$

Substituting Eq. (A.49) into Eq. (A.48), it is straightforward to compute Eq. (3.27). ∎

A.11 Proof of Proposition 3.3.2

From the definitions of $\alpha(\chi;\eta)$, $\beta(\chi;\eta)$ and inspection of Eq. (3.27), it follows that $A_L(1) = A_L(0)$. Furthermore, if $\chi \in (0,1)$, the sign of $A_L(\chi) - A_L(0)$ is equivalent to the sign of

$$\delta(\chi;\eta) \equiv \frac{1}{1+\theta(\bar{\zeta}-1)}\left(1-\chi^{\frac{k-(1-\theta)\bar{\xi}}{k}}\right)\left[\eta\left(1+\chi^{\frac{\bar{\xi}}{k}}\right)-1\right] - \left(1-\chi^{\frac{k-\bar{\xi}}{k}}\right)\chi^{\frac{\theta\bar{\xi}}{k}}\left[\eta\left(1+\chi^{\frac{\bar{\xi}}{k}}\right)^{1-\theta}-1\right]. \tag{A.50}$$

Noting further that

$$\frac{1}{1+\theta(\bar{\zeta}-1)}\left(1-\chi^{\frac{k-(1-\theta)\bar{\xi}}{k}}\right) > \left(1-\chi^{\frac{k-\bar{\xi}}{k}}\right)\chi^{\frac{\theta\bar{\xi}}{k}} \tag{A.51}$$

holds for any possible $\chi \in (0,1)$, it is straightforward to show that $\delta(\chi;\eta) > 0$ must hold if $\eta(1+\chi^{\bar{\xi}/k}) \geq 1$, or, equivalently, if $\chi \geq [(1-\eta)/\eta]^{k/\bar{\xi}}$.

But what is the sign of $\delta(\chi;\eta)$ if $\chi < \bar{\chi} \equiv [(1-\eta)/\eta]^{k/\bar{\xi}}$, where $\bar{\chi} < 1$ follows from $\eta > 0.5$? To answer this question, we can first note that if $\chi < \bar{\chi}$, condition $\delta(\chi;\eta) >,=,< 0$ is equivalent to condition $\delta_0(\chi;\eta) >,=,< \delta_1(\chi)$, with

$$\delta_0(\chi;\eta) \equiv \frac{1-\eta\left(1+\chi^{\frac{\bar{\xi}}{k}}\right)^{1-\theta}}{1-\eta\left(1+\chi^{\frac{\bar{\xi}}{k}}\right)}, \quad \delta_1(\chi) \equiv \frac{1}{1+\theta(\bar{\zeta}-1)}\frac{1-\chi^{\frac{k-(1-\theta)\bar{\xi}}{k}}}{\left(1-\chi^{\frac{k-\bar{\xi}}{k}}\right)\chi^{\frac{\theta\bar{\xi}}{k}}}. \tag{A.52}$$

It is easily confirmed that $\delta_0(\chi;\eta)$ increases in χ over the relevant interval, reaching a minimum function value of $\delta_0(0;\eta) = 1$ at $\chi = 0$. Accordingly, $\delta_0(\chi;\eta)$ reaches a maximum function value of ∞ at $\bar{\chi}$. In a similar way, we can show that $\delta_1(\chi)$ is decreasing in χ, reaching a maximum function value of ∞ at $\chi = 0$ and a minimum function value of 1 at $\chi = 1$. Putting together, this implies that there exists a unique $\hat{\chi}(\eta)$ such that $\delta_0(\chi;\eta) >,=,< \delta_1(\chi)$ and thus $\delta(\chi;\eta) >,=,< 0$ if $\chi >,=,< \hat{\chi}(\eta)$. Finally, accounting for $\partial\delta_0(\chi;\eta)/\partial\eta > 0$, it follows that $\hat{\chi}(\eta)$ falls in η and reaches a minimum value of 0 at $\eta = 1$. In this case $\delta(\chi;0) > 0$ holds for any $\chi \in (0,1)$. Furthermore, $\hat{\chi}(\eta)$ reaches a maximum value of 1 at $\eta = 0$, implying that in this case $\delta(\chi;0) < 0$ must hold for any $\chi \in (0,1)$. This completes the formal discussion of the properties of $A_L(\chi)$. ∎

A.12 Income inequality among employed and unemployed workers

To characterise income inequality among *all* production workers, we rely on the Lorenz curve for labour income. Since this Lorenz curve now also captures unemployed individuals, it consists of three segments. The first segment represents the share of income attributed to those who do not have a job. Abstracting from unemployment compensation, it is clear that the income share of this group is zero, and we can thus note that the respective Lorenz curve segment is given by $Q_U^0(\lambda) = 0$ and relevant for all $\lambda \in [0, U)$.

To calculate the second segment of the Lorenz curve, we follow the steps in Appendix A.10 and combine the labour income share of workers employed in purely domestic firms with a productivity level up to $\bar{\varphi} \in [\varphi^d, \varphi^o)$ – as determined by Eq. (A.35) – with the share of *all* production workers who are either unemployed or employed in firms up to productivity $\bar{\varphi}$:

$$\lambda = U + \frac{1-U}{\beta(\chi;\eta)} \left[1 - \left(\frac{\bar{\varphi}}{\varphi^d}\right)^{(1-\theta)\bar{\xi}-k}\right]. \tag{A.53}$$

This gives the second segment of the Lorenz curve for the distribution of labour income

$$Q_U^1(\lambda) = \frac{1 - \left[1 - \beta(\chi;\eta)\frac{\lambda-U}{1-U}\right]^{\frac{k-\bar{\xi}}{k-(1-\theta)\bar{\xi}}}}{\alpha(\chi;\eta)}, \tag{A.54}$$

which is relevant for $\lambda \in [U, b_U)$, with $b_U \equiv U + (1-U)(1 - \chi^{1-(1-\theta)\bar{\xi}/k})/\beta(\chi;\eta)$.

To determine the third segment of the Lorenz curve, we compute the share of total domestic labour income accruing to workers who are either unemployed or employed in firms with a productivity level up to $\bar{\varphi} \in [\varphi^o, \infty)$ – as represented by Eq. (A.39) – with the share of production workers who are either unemployed or employed by these firms:

$$\lambda = 1 - \eta\left(1 + \chi^{\bar{\xi}/k}\right)^{1-\theta} \frac{1-U}{\beta(\chi;\eta)} \left(\frac{\bar{\varphi}}{\varphi^d}\right)^{(1-\theta)\bar{\xi}-k}. \tag{A.55}$$

This allows us to calculate the third segment of the Lorenz curve

$$Q_U^2(\lambda) = 1 - \frac{\eta\left(1+\chi^{\bar{\xi}/k}\right)}{\alpha(\chi;\eta)} \left[\left(\frac{1-\lambda}{1-U}\right) \frac{\beta(\chi;\eta)}{\eta\left(1+\chi^{\bar{\xi}/k}\right)^{1-\theta}}\right]^{\frac{k-\bar{\xi}}{k-(1-\theta)\bar{\xi}}}, \tag{A.56}$$

which is relevant if $\lambda > b_U$. Putting the three segments together, gives the (extended) Lorenz curve for labour income distribution

$$Q_U(\lambda) \equiv \begin{cases} 0 & \text{if} \quad \lambda \in [0, U) \\ Q_U^1(\lambda) & \text{if} \quad \lambda \in [U, b_U) \\ Q_L^2(\lambda) & \text{if} \quad \lambda \in [b_U, 1] \end{cases} . \tag{A.57}$$

The Gini coefficient for the distribution of labour income can then be computed according to $A_U(\chi) \equiv 1 - 2\int_0^1 Q_U(\lambda)d\lambda$, and is given by the respective expression in Footnote 32. ■

A.13 Derivation of the Theil index in Eq. (3.30)

Applying the decomposition rule in Eq. (3.29), we can write

$$T = a_S T_S + a_U T_U + a_S \ln\left(a_S \frac{N}{N-L}\right) + a_U \ln\left(a_U \frac{N}{L}\right), \tag{A.58}$$

where a_S, a_U are the income shares of self-employed agents and production workers, respectively, given by Eq. (3.31). Accounting for Eq. (3.15) and considering $\sigma - 1 = \rho/(1-\rho)$, we can furthermore compute

$$a_S \frac{N}{N-L} = \frac{\bar{\zeta}\rho\gamma + 1 - \rho}{\rho\gamma + 1 - \rho}, \quad a_U \frac{N}{L} = \frac{\bar{\zeta}\rho\gamma + 1 - \rho}{\bar{\zeta}(\rho\gamma + 1 - \rho)}. \tag{A.59}$$

Substitution of Eq. (A.59) into Eq. (A.58) allows us to write

$$T = a_S T_S + a_L T_U + (a_S + a_U) \ln\left(\frac{\bar{\zeta}\rho\gamma + 1 - \rho}{\bar{\zeta}(\rho\gamma + 1 - \rho)}\right) + a_S \ln\left(\bar{\zeta}\right), \tag{A.60}$$

Finally, noting that $a_S + a_U = 1$ and $a_S/\bar{\zeta} + a_U = [\bar{\zeta}\rho\gamma + 1 - \rho]/[\bar{\zeta}(\rho\gamma + 1 - \rho)]$ hold, according to Eq. (3.31), we obtain Eq. (3.30). ■

A.14 The concept of Lorenz dominance

We now consider a second criterion for ranking distributions and look at the criterion of Lorenz dominance. Thereby, we say that distribution A Lorenz dominates distribution B if the Lorenz curve of A lies above the Lorenz curve of B for any cumulative share of the population. Since the Lorenz dominance is equivalent to mean-preserving second-order stochastic dominance, all measures of inequality that respect this criterion – such as the Gini coefficient or the Theil index – rank A as a more equal distribution than B if A Lorenz dominates B.

The Lorenz curve for the income distribution of self-employed agents under autarky is given by

$$Q_S^a(\mu) = 1 - (1-\mu)^{\frac{k-\bar{\xi}}{k}}. \tag{A.61}$$

Hence, the income distribution of self-employed agents under autarky Lorenz dominates the respective income distribution under partial offshoring if $Q_S(\mu) < Q_S^a(\mu)$ holds for any $\mu \in (0,1)$. We have to check this inequality separately for the three segments of $Q_S(\mu)$. Let us first look at domain $\mu \in \big(0, \chi/(1+\chi)\big)$. In this case, $Q_S(\mu) < Q_S^a(\mu)$, is equivalent to $D_S^0(\mu, b) \equiv b\mu - 1 + (1-\mu)^b < 0$, with $b \equiv 1/\zeta$. Twice differentiating $D_S^0(\mu, b)$ with respect to b gives

$$\frac{\partial D_S^0(\mu,b)}{\partial b} = \mu + \ln(1-\mu)(1-\mu)^b, \qquad \frac{\partial^2 D_S^0(\mu,b)}{\partial b^2} = \left[\ln(1-\mu)\right]^2 (1-\mu)^b. \tag{A.62}$$

with $\partial D_S^0(\mu,0)/\partial b = \mu + \ln(1-\mu) < 0$, $\partial D_S^0(\mu,1)/\partial b = \mu + \ln(1-\mu)(1-\mu) > 0$, and $\partial^2 D_S^1(\mu,b)/\partial b^2 > 0$. Accounting for $D_S^0(\mu,0) = D_S^0(\mu,1) = 0$, we can therefore conclude that $D_S^0(\mu,b) < 0$ and thus $Q_S(\mu) < Q_S^a(\mu)$ must hold in the relevant parameter range.

For domain $\mu \in \big[\chi/(1+\chi), 1/(1+\chi)\big)$, it follows from Eqs. (A.17) and (A.61) that $Q_S(\mu) < Q_S^a(\mu)$ is equivalent to $D_S^1(\mu,b) \equiv (b-1)\chi + [1-(1+\chi)^{b-1}](1+\chi)(1-\mu)^b < 0$. Therefore, $\partial D_S^1(\mu,b)/\partial\mu < 0$ implies that $D_S^1(\chi/(1+\chi), b) \equiv \hat{D}_S^1(b) = (b-1)\chi + (1+\chi)^{1-b} - 1 < 0$ is sufficient for $Q_S^1(\mu) < Q_S^a(\mu)$ to hold in the relevant parameter domain. Twice differentiating $\hat{D}_S^1(b)$ yields $d\hat{D}_S^1(b)/db = \chi - \ln(1+\chi)(1+\chi)^{1-b}$, $d^2\hat{D}_S^1(b)/db = \left[\ln(1+\chi)\right]^2 (1+\chi)^{1-b}$. Accounting for $d\hat{D}_S^1(0)/db = \chi - \ln(1+\chi)(1+\chi) < 0$, $d\hat{D}_S^1(1)/db = \chi - \ln(1+\chi) > 0$, and $d^2\hat{D}_S^1(b)/db^2 > 0$, it follows from $\hat{D}_S^1(0) = \hat{D}_S^1(1) = 0$ that $\hat{D}_S^1(b) < 0$ and thus $Q_S(\mu) < Q_S^a(\mu)$ must hold for all $\mu \in \big[\chi/(1+\chi), 1/(1+\chi)\big)$.

Finally, we look at domain $\mu \in \big[1/(1+\chi), 1\big]$. In this case, $Q_S(\mu) < Q_S^a(\mu)$ is equivalent to $D_S^2(\mu,b) \equiv -(1+\chi)^b(1-\mu)^b + b(1+\chi)(1-\mu) < 0$, according to Eqs. (A.18) and (A.61). Twice differentiating $D_S^2(\mu,b)$ with respect to μ gives $\partial D_S^2(\mu,b)/\partial\mu = b(1+\chi)^b(1-\mu)^{b-1}[1-(1+\chi)^{1-b}(1-\mu)^{1-b}]$, $\partial^2 D_S^2(\mu,b)/\partial\mu^2 = b(1-b)(1+\chi)^b(1-\mu)^{b-2} > 0$. Accounting for $\partial D_S^2(1/(1+\chi),b)/\partial\mu = b(1+\chi)(\chi^{b-1}-1) > 0$, it is thus immediate that $D_S^2(1,b) = 0$ is sufficient for $Q_S(\mu) < Q_S^a(\mu)$ to hold in the relevant parameter domain.

Putting together, we can thus conclude that $Q_S(\mu) < Q_S^a(\mu)$ holds for any $\mu \in (0,1)$, which proves that the income distribution of self-employed agents under autarky Lorenz dominates the respective income distribution in an offshoring equilibrium for arbitrary values of $\chi \in (0,1)$.

The Lorenz curve for the distribution of labour income under autarky has two segments and is given by

$$Q_U^a(\lambda) = \begin{cases} 0 & \text{if} \quad \lambda \in [0, U^a) \\ 1 - \left(\frac{1-\lambda}{1-U^a}\right)^{\frac{k-\bar{\xi}}{k-(1-\theta)\bar{\xi}}} & \text{if} \quad \lambda \in [U^a, 1] \end{cases}, \tag{A.63}$$

where $U^a = \theta(\bar{\zeta}-1)/[1+\theta(\bar{\zeta}-1)]$, according to Eq. (3.24). The ranking of $Q_U^a(\lambda)$ and $Q_U(\lambda)$ depends on the unemployment rate of production workers in the offshoring scenario relative to autarky. Furthermore, as outlined in the main text, the ranking of $U >,=,< U^a$ is equivalent to the ranking of $1 >,=,< \Delta(\chi;\eta)$ and thus equivalent to the ranking of $\alpha(\chi;\eta) >,=,< \beta(\chi;\eta)$. From Appendix A.9 we know that there exists a unique $\hat{\chi} \in (0,1)$, such that $\alpha(\chi;\eta) >,=,< \beta(\chi;\eta)$ if $\chi >,=,< \hat{\chi}$.

Let us first consider $\chi \geq \hat{\chi}$, which corresponds to an offshoring equilibrium with $U \geq U^a$. In this case, we have $Q_U^a(\lambda) = Q_U(\lambda) = 0$ for all $\lambda \in [0, U^a)$ and $Q_U^a(\lambda) > Q_U(\lambda) = 0$ for all $\lambda \in [U^a, U)$. Furthermore, combining Eqs. (A.54) and (A.63), it follows that, for domain $\lambda \in [U, b_U)$, the ranking of $Q_U(\lambda) >,=,< Q_U^a(\lambda)$ is equivalent to the ranking of

$$D_U^1(\hat{\lambda}) \equiv 1 - \alpha(\chi;\eta) + \alpha(\chi;\eta)\left(1-\hat{\lambda}\right)^{\hat{b}} \Delta(\chi;\eta)^{\hat{b}} - \left[1-\beta(\chi;\eta)\hat{\lambda}\right]^{\hat{b}} >,=,< 0, \tag{A.64}$$

where $\hat{b} \equiv 1/[1+\theta(\bar{\zeta}-1)]$ and $\hat{\lambda} \equiv (\lambda - U)/(1-U)$. Differentiating $D_U^1(\hat{\lambda})$ gives

$$\frac{dD_U^1(\hat{\lambda})}{d\hat{\lambda}} = \frac{\hat{b}\alpha(\chi;\eta)\Delta(\chi;\eta)^{\hat{b}}}{\left(1-\hat{\lambda}\right)^{1-\hat{b}}} \left[\Delta(\chi;\eta)^{1-\hat{b}} \left(\frac{1-\hat{\lambda}}{1-\beta(\chi;\eta)\hat{\lambda}}\right)^{1-\hat{b}} - 1\right]. \tag{A.65}$$

Consider first the case of $\beta(\chi;\eta) \leq 1$. Since $\chi \geq \hat{\chi}$ implies $\beta(\chi;\eta) \leq \alpha(\chi;\eta)$ and thus $\Delta(\chi;\eta) \leq 1$, it is immediate that $\beta(\chi;\eta) \leq 1$ is sufficient for $dD_U^1(\hat{\lambda})/d\hat{\lambda} < 0$. Noting further that $\lambda = U$ implies $\hat{\lambda} = 0$ and thus $D_U^1(0) = \alpha(\chi;\eta)[\Delta(\chi;\eta)^{\hat{b}} - 1] < 0$, we can therefore safely conclude that $Q_U(\lambda) < Q_U^a(\lambda)$ holds for all $\lambda \in [U, b_U)$ in this case.

But what happens if $\beta(\chi;\eta) > 1$? In this case, we cannot rule out that $dD_U^1(\hat{\lambda})/d\hat{\lambda} > 0$.

However, computing the second derivative of $D_U^1(\hat{\lambda})$, we obtain

$$\frac{d^2 D_U^1(\hat{\lambda})}{d\hat{\lambda}^2} = \frac{1-\hat{b}}{1-\hat{\lambda}} \left\{ \frac{dD_U^1(\hat{\lambda})}{d\hat{\lambda}} - \frac{\hat{b}\alpha(\chi;\eta)\Delta(\chi;\eta)}{\left(1-\hat{\lambda}\right)^{1-\hat{b}}} \left(\frac{1-\hat{\lambda}}{1-\beta(\chi;\eta)\hat{\lambda}} \right)^{1-\hat{b}} \frac{1-\beta(\chi;\eta)}{1-\beta(\chi;\eta)\hat{\lambda}} \right\}. \tag{A.66}$$

From inspection of Eqs. (A.65) and (A.66) we can therefore conclude that $dD_U^1(\hat{\lambda})/d\hat{\lambda} \geq 0$ is sufficient for $d^2 D_U^1(\hat{\lambda})/d\hat{\lambda}^2 > 0$ if $\beta(\chi;\eta) > 1$. To see this, note that $\beta(\chi;\eta)\hat{\lambda} < \beta(\chi;\eta)\hat{\lambda}_U$, with $\hat{\lambda}_U \equiv (b_U - U)/(1-\lambda)$ must hold on the relevant parameter domain. Substituting for b_U, we obtain $\beta(\chi;\eta)\hat{\lambda} < 1 - \chi^{1-(1-\theta)\bar{\xi}/k} < 1$. From inspection of Eqs. (A.65) and (A.66) it therefore follows that if $dD_U^1(\hat{\lambda})/d\hat{\lambda} \geq 0$ holds for some $\hat{\lambda}_0 \in (0, \hat{\lambda}_U)$, then $dD_U^1(\hat{\lambda})/d\hat{\lambda} > 0$ must hold for all $\hat{\lambda} \in (\hat{\lambda}_0, \hat{\lambda}_U)$. Furthermore, recollecting from above that $D_U^1(0) < 0$, this implies that if $D_U^1(\hat{\lambda}) \geq 0$ holds for some $\hat{\lambda} \in (0, \hat{\lambda}_U)$, then $D_U^1(\hat{\lambda}_U) > 0$ must hold as well. Accordingly, we can infer insights on the sign of $D_U^1(\hat{\lambda})$ by evaluating Eq. (A.64) at $\hat{\lambda} = \hat{\lambda}_U$. This gives

$$D_U^1(\hat{\lambda}_U) = \alpha(\chi;\eta)^{1-\hat{b}} \chi^{1-\bar{\xi}/k} \left[\eta\left(1+\chi^{\bar{\xi}/k}\right)\right]^{\hat{b}} \left[\left(1+\chi^{\bar{\xi}/k}\right)^{-\theta\hat{b}} - \left(\frac{\eta\left(1+\chi^{\bar{\xi}/k}\right)}{\alpha(\chi;\eta)} \right)^{1-\hat{b}} \right]. \tag{A.67}$$

Since $\beta(\chi;\eta) > 1$ implies $\alpha(\chi;\eta) > 1$ if $\chi \geq \hat{\chi}$, it is immediate that $\alpha(\chi;\eta) < \eta(1+\chi^{\bar{\xi}/k})$, and this implies $D_U^1(\hat{\lambda}_U) < 0$. Putting together, we can therefore safely conclude that $Q_U(\lambda) < Q_U^a(\lambda)$ holds for all $\lambda \in [U, b_U)$ irrespective of the ranking of $\beta(\chi;\eta) >,=,< 1$.

In a final step, we have to look at domain $\lambda \in [b_U, 1]$. According to Eqs. (A.56) and (A.63), for this parameter domain the ranking of $Q_U(\lambda) >,=,< Q_U^a(\lambda)$ is equivalent to the ranking of

$$D_U^2(\hat{\lambda}) \equiv \left[1 - \left(1+\chi^{\bar{\xi}/k}\right)^{\theta\hat{b}} \left(\frac{\eta\left(1+\chi^{\bar{\xi}/k}\right)}{\alpha(\chi;\eta)} \right)^{1-\hat{b}} \right] \left(1-\hat{\lambda}\right)^{\hat{b}} \Delta(\chi;\eta)^{\hat{b}} >,=,< 0. \tag{A.68}$$

Notably, the sign of $D_U^2(\hat{\lambda})$ does not depend on the specific level of $\hat{\lambda}$, so that $\text{sgn}[D_U^2(\hat{\lambda})] = \text{sgn}[D_U^2(\hat{\lambda}_U)]$. However, since $D_U^1(\hat{\lambda}_U) = D_U^2(\hat{\lambda}_U)$ holds by definition, it follows that $Q_U(\lambda) < Q_U^a(\lambda)$ extends to interval $\lambda \in [b_U, 1]$. Summing up, we can thus conclude that the income distribution of production workers under autarky Lorenz dominates the income distribution of production workers in the offshoring equilibrium if the share of offshoring firms is sufficiently high, i.e. if $\chi \geq \hat{\chi}$.

Let us now consider $\chi < \hat{\chi}$, which implies $\Delta(\chi;\eta) > 1$ and thus $U < U^a$. In this case, we have $Q_U(\lambda) = Q_U^a(\lambda) = 0$ for all $\lambda \in [0, U]$ and $Q_U(\lambda) > Q_U^a(\lambda) = 0$ for all $\lambda \in (U, U^a)$. For domain $\lambda \in (U, b_U)$, the ranking of $Q_U(\lambda) >, =, < Q_U^a(\lambda)$ depends on the sign of $D_U^1(\hat{\lambda})$, where $D_U^1(0) = \alpha(\chi;\eta)[\Delta(\chi;\eta)^{\hat{b}} - 1] > 0$ holds if $\chi < \hat{\chi}$. But what can we say about the sign of $D_U^1(\hat{\lambda})$ if $\hat{\lambda} > U$? To answer this question, it is worth looking at Eq. (A.66). From the formal discussion in Appendix A.9, we know that $\Delta(\chi;\eta) > 1$ requires $\hat{\Omega}(\bar{\eta}) \equiv (\bar{\eta}\vartheta^{-\theta} - 1)(\vartheta - 1)^{\theta} - \bar{\eta} + 1 > 0$, where $\vartheta \equiv 1 + \chi^{\bar{\xi}/k}$ and $\bar{\eta} \equiv \eta\vartheta$. In view of $\hat{\Omega}'(\bar{\eta}) = (1 - 1/\vartheta)^{\theta} - 1 < 0$ and $\hat{\Omega}(1) = (\vartheta^{-\theta} - 1)(\vartheta - 1) < 0$, we can conclude that $\bar{\eta} = \eta(1 + \chi^{\bar{\xi}/k}) < 1$ is necessary for $\Delta(\chi;\eta) > 1$. This implies that $\beta(\chi;\eta) < 1$ must hold for all $\chi < \hat{\chi}$. Hence, if $D_U^1(\hat{\lambda})$ has an extremum at $\hat{\lambda} \in (0, \hat{\lambda}_U)$, this extremum must be a maximum. In view of $D_U^1(0) > 0$, we can therefore conclude that $D_U^1(\hat{\lambda})$ is positive for all $\lambda \in [U, b_U)$ if $D_U^1(\hat{\lambda}_U) \geq 0$, while $D_U^1(\hat{\lambda}_U) < 0$ implies that there exists a unique $\lambda_0 \in [U, b_U)$ such that $D_U^1(\hat{\lambda}) >, =, < 0$ if $\lambda_0 >, =, < \lambda$. Noting finally that $\text{sgn}[D_U^1(\hat{\lambda}_U)] = \text{sgn}[D_U^2(\hat{\lambda})]$ holds for all $\lambda \in [b_U, 1)$ and accounting for $\lim_{\chi \to 0} D_U^2(\hat{\lambda}_U) = (1 - \eta^{1-\hat{b}})(1 - \hat{\lambda})^{\hat{b}}\Delta(\chi;\eta)^{\hat{b}} > 0$, $\lim_{\chi \to \hat{\chi}} D_U^2(\hat{\lambda}_U;\beta) < 0$ (see our extensive discussion for domain $\chi \geq \hat{\chi}$), the following conclusion is immediate: For sufficiently small χ, the distribution of labour income with offshoring Lorenz dominates the respective distribution without offshoring. For χ smaller than but close to $\hat{\chi}$, Lorenz curves Q_U^a and Q_U intersect and it is therefore not possible to rank the distributions of labour income with and without offshoring according to the criterion of Lorenz dominance. This completes our discussion on Lorenz curve dominance. ■

A.15 Economy-wide income distribution if $\theta > 0$

In the subsequent analysis it is useful to introduce the Theil index for the income distribution within the group of employed production workers, which we denote by T_L. Thereby, T_L is linked to T_U according to $T_U = T_L - \ln(1 - U)$. This allows us to rewrite Eq. (A.58) as follows:

$$T = a_S T_S + a_U T_L + a_S \ln\left(a_S \frac{N}{N - L}\right) + a_U \ln\left(a_U \frac{N}{L(1 - U)}\right), \tag{A.69}$$

Following the analysis in the main text step by step and substituting Eq. (3.24) for U, we thus obtain

$$T = \frac{1-\rho}{\rho\gamma+1-\rho} T_S + \frac{\rho\gamma}{\rho\gamma+1-\rho} T_L + \ln\left(\frac{\bar{\zeta}\rho\gamma+1-\rho}{\bar{\zeta}(\rho\gamma+1-\rho)}\right)$$
$$+ \frac{1-\rho}{\rho\gamma+1-\rho}\ln\bar{\zeta} - \frac{\rho\gamma}{\rho\gamma+1-\rho}\ln\left(\frac{\Delta(\chi;\eta)}{1+\theta(\bar{\zeta}-1)}\right). \tag{A.70}$$

In autarky, we can explicitly compute the Theil indices for the income distribution of self-employed agents and production workers, respectively:

$$T_S^a = \frac{1}{\bar{\zeta}-1}\int_1^\infty x^{-\frac{k}{\xi}}\left[\ln x - \ln\bar{\zeta}\right]dx = \bar{\zeta} - 1 - \ln\bar{\zeta}, \tag{A.71}$$

and

$$T_L^a = \frac{1}{1+\theta(\bar{\zeta}-1)}\int_1^\infty y^{-\frac{1+\theta(\bar{\zeta}-1)}{\theta(\bar{\zeta}-1)}}\left\{\ln y - \ln\left[1+\theta(\bar{\zeta}-1)\right]\right\}dy$$
$$= \theta(\bar{\zeta}-1) - \ln\left[1+\theta(\bar{\zeta}-1)\right]. \tag{A.72}$$

Substituting for T_S, T_L and setting $\chi = 0$ then yields

$$T^a = [1-\rho(1-\theta)](\bar{\zeta}-1] + \ln\left(\frac{1+\rho(\bar{\zeta}-1)}{\bar{\zeta}}\right). \tag{A.73}$$

While we are not able to rank T and T^a for arbitrary levels of χ, we can at least compare Theil indices for the two limiting cases $\chi = 0$ and $\chi = 1$. Since $T_L = T_L^a$ and $T_S > T_S^a$ hold if $\chi = 1$, we can safely conclude that $T - T_a > \hat{\Delta}_T(\theta;\hat{a})$, with

$$\hat{\Delta}_T(\theta;\hat{a}) \equiv \underbrace{\frac{\rho(1-\rho)(1-\eta)}{\rho\eta+1-\rho}\frac{\hat{a}}{1-\hat{a}} + \ln\left(\frac{\rho\eta+(1-\rho)(1-\hat{a})}{[\rho\eta+1-\rho]\left[1-(1-\rho)\hat{a}\right]}\right)}_{\hat{\Delta}_T^1(\hat{a})}$$
$$+ \underbrace{\frac{\rho\theta}{\rho\eta+1-\rho}\left[\eta\ln 2 - \frac{(1-\rho)(1-\eta)\hat{a}\theta}{1-\hat{a}}\right]}_{\hat{\Delta}_T^2(\theta;\hat{a})} \tag{A.74}$$

and $\hat{a} \equiv 1 - 1/\bar{\zeta}$. Differentiating $\Delta_T^1(\hat{a})$ gives

$$\frac{d\hat{\Delta}_T^1(\hat{a})}{d\hat{a}} = \frac{\rho^2(1-\rho)(1-\eta)\hat{a}\left[\rho\eta+(1-\rho)(1-\hat{a})+\eta(1-\hat{a})\right]}{[\rho\eta+1-\rho]\left[\rho\eta+(1-\rho)(1-\hat{a})\right]\left[1-(1-\rho)\hat{a}\right](1-\hat{a})^2} > 0. \tag{A.75}$$

In view of $\hat{\Delta}_T^1(0) = 0$, this implies that $\hat{\Delta}_T^1(\hat{a}) > 0$ holds for all $\hat{a} \in (0,1)$. While the sign of $\hat{\Delta}_T^2(\theta;\hat{a})$ is not clear in general, it is immediate that $\hat{\Delta}_T(\theta;\hat{a}) > 0$ holds for sufficiently small levels of θ. This completes our discussion on the Theil index. ∎

A.16 Source code for the calibration exercises in Section 3.5

The calibration exercise has been executed in *Mathematica*.[7] We offer here the source code as well as the parameter estimates used in our calibration. At first, we set parameter values: $k = 4.306$, $\sigma = 6.698$ and $\theta = 0.102$, based on the results in Egger, Egger, and Kreickemeier (2013), and $\eta = 0.75$, based on Blinder (2009) and Blinder and Krueger (2013).

```
k=4.306;
σ=6.698;
θ=0.102;
η=0.75;
```

Furthermore, regarding the extent of external increasing returns to scale, we consider the two polar cases $\varepsilon = 0$ and $\varepsilon = 1$. In addition, we account for $\varepsilon = 0.56$ as reported by Ardelean (2011).

```
ε={0,0.56,1}
```

As all variables of interest can be expressed in terms of the share of offshoring firms, χ, we define

```
χ=.;
χG=1265/8466;
```

where χG is the share of offshoring firms in Germany as reported by Moser, Urban, and Weder di Mauro (2009). We then define ρ, ξ and ζ and check that $k > \xi$ holds.[8]

```
ρ=(σ-1)/σ;
ξ=(σ-1)/(1+θ(σ-1));
ζ=k/(k-ξ);
If[k<=ξ, Print["Error:k<=ξ"]];
```

[7] A self-contained Computable-Data-File (CDF), which can be run on the free to use CDF-player offered by Wolfram Research, Inc. under `http://www.wolfram.com/cdf-player/`, can be obtained from the authors upon request.

[8] In the source code, we use ξ instead of $\bar{\xi}$ and ζ instead of $\bar{\zeta}$ to save on notation.

We also define $\alpha(\chi;\eta)$ and $\beta(\chi;\eta)$ from Eq. (3.25) as well as $\gamma(\chi;\eta)$ and $\Delta(\chi,\eta)$ as specified in the main text.

```
α=1+χ^((k-ξ)/k)(η(1+χ^(ξ/k))-1);
β=1+χ^((k-(1-θ)ξ)/k)(η(1+χ^(ξ/k))^(1-θ)-1);
γ=(1+η*χ-(1-η)χ^(1-ξ/k))/(1+χ);
Δ=β/α;
```

Now we can turn to aggregate income in the source country relative to autarky, see the proof of Proposition 3.2.3:

```
T1=(1+γ(σ-1))(1+ζ(σ-1))/(σ(1+γ*ζ(σ-1)));
T2=((1+χ)(1+γ*ζ(σ-1))/(1+ζ(σ-1)))^((σ-1-ε*k)/(k(σ-1)));
T3=(1+χ)^(1/(σ-1));
Φ=T1*T2*T3;
```

Eq. (3.26) establishes

```
Λ=((θ(ζ-1)+1-Δ)/(θ(ζ-1)))*((1+ζ(σ-1))γ)/(1+ζ(σ-1)γ);
```

where $u/u^a = \Lambda$ and

```
ua=(θ(ζ-1)/(1+θ(ζ-1))*ζ(σ-1)/(1+ζ(σ-1));
```

Finally, inter-group inequality between entrepreneurs and workers as well as intra-group inequality within both groups, each normalised to one for its respective autarky level, follow from Eqs. (3.20), (3.21) and (3.27).

```
ω=1+(1-1/ζ)χ;
AM=1+χ(2-χ)/(ζ+(ζ-1)χ);
AL=1+(2(1-χ^((k-(1-θ)ξ)/k))(α-1)-2(1+θ(ζ-1))(1-χ^((k-ξ)/k))(β-1))/(α*β*θ(ζ-1));
```

We now turn to the determination of the Theil indices. Therefore, we first need to specify the income share of entrepreneurs, freelance offshoring workers and production workers. This gives

```
aM=(1-ρ)/(ζ(ρ*γ+1-ρ))*((ζ+χ(ζ-1))/(1+χ);
aF=(1-ρ)/(ζ(ρ*γ+1-ρ))*(χ/(1+χ));
aU=(ρ*γ)/(ρ*γ+1-ρ);
```

respectively. Furthermore, we also determine the income share of self-employed agents, as defined in Eq. (3.31):

```
aS=(1-ρ)/(ρ*γ+1-ρ);
```

Average income of the three subgroups – entrepreneurs, agents in the offshoring service sector, and production workers – relative to the economy-wide income average is given by

```
vM=((ζ*ρ*γ+1-ρ)/(ρ*γ+1-ρ))*(1+(1-1/ζ)χ);
vF=(ζ*ρ*γ+1-ρ)/(ζ*(ρ*γ+1-ρ));
vU=(ζ*ρ*γ+1-ρ)/(ζ*(ρ*γ+1-ρ));
```

while for the self-employed we obtain

```
vS=(ζ*ρ*γ+1-ρ)/(ρ*γ+1-ρ);
```

We now determine the product of income ratios and log income ratios for entrepreneurial income multiplied by the relative frequency the respective income ratios are realized. For purely domestic firms, this gives

```
Alt1=(k*x^(ξ-k-1)/(ζ+χ(ζ-1)))*Log[(x^ξ)/(ζ+χ(ζ-1))];
```

while for offshoring firms, we obtain

```
Alt2=((k*((1+χ^(ξ/k))*x^ξ-1)x^(-k-1))/(ζ+χ(ζ-1)))*Log[(((1+χ^(ξ/k))*x^ξ)
-1)/(ζ+χ(ζ-1))];
```

We can compute similar expressions for production workers and obtain

```
Alt3=(Δ/(θ(ζ-1)β))y^(((1-θ)*ξ-k)/(θ*ξ))*Log[y*Δ/(1+θ(ζ-1))];
```

for workers employed in purely domestic firms and

```
Alt4=(Δ/(θ(ζ-1)β))*η*((1+χ^(ξ/k))^(k/ξ))y^(((1-θ)*ξ-k)/(θ*ξ))*Log[y*Δ/(1
+θ(ζ-1))];
```

for workers employed in exporting firms.

In a last step we evaluate the above defined functions at $\chi = 0.001, 0.01, 0.1, 0.25, 0.5, 0.75, 0.9$ and $\chi = \chi_G$ to produce the results in Tables 3.1 and 3.2.

```
Do[z=z0;
```

We start with the two Gini coefficients and the measure for inter-group inequality and evaluate

```
nω=ω/.{χ->z};
nAM=AM /.{χ->z};
nAL=AL /.{χ->z};
```

We now turn to the Theil index for entrepreneurial income, which in autarky can be computed according to

```
Alt00=Alt1/.{χ->0};
TMa= NIntegrate[Alt00, {x,1,Infinity}];
```

The respective Theil index in the case of offshoring can be determined according to

```
Alt11=Alt1/.{χ->z};
Alt22=Alt2/.{χ->z};
TM=NIntegrate[Alt11, {x,1,z^(-1/k)}]+NIntegrate[Alt22, {x,z^(-1/k),Infinity}];
```

In a similar vein, we can compute the Theil index for income of employed production workers under autarky and in the scenario with offshoring. This gives

```
Alt55=Alt3/.{χ->0};
TLa=NIntegrate[Alt55, {y,1,Infinity}];
```

and

```
Alt33=Alt3/.{χ->z};
Alt44=Alt4/.{χ->z};
TL=NIntegrate[Alt33, {y,1,z^(-θ*ξ/k)}]+NIntegrate[Alt44,
{y,z^(-θ*ξ/k)(1+z^(ξ/k))^θ, Infinity}];
```

respectively. Thereby, it is notable that in the scenario with offshoring, firms which shift production abroad pay a wage premium to their domestic workers, and this wage premium is captured by an upward shift of the lower bound of the second integral in the equation for `TL`.

The economy-wide Theil index under autarky is then given by

```
Ta1=aM*TMa+aU*TLa+aM*Log[vM]+aF*Log[vF]+aU*Log[vU]-aU*Log[Δ/(1+θ(ζ-1)];
Ta=Ta1/.{χ->0};
```

where $1 - U^a = \Delta^a/[1+\theta(\zeta-1)]$ has been considered, according to (3.24). The economy-wide Theil index in the scenario with offshoring is given by

```
T1=aM*TM+aU*TL+aM*Log[vM]+aF*Log[vF]+aU*Log[vU]-aU*Log[Δ/(1+θ(ζ-1)];
T=T1/.{χ->z};
```

To avoid rounding errors, we can manipulate the result in the following way

```
If[TL<TLa + 0.1^(10)&&TL>TLa - 0.1^(10), TL=TLa];
```

Finally, we can compute T_U, considering the calibrated values of T_L. Accounting for

```
Δa=Δ/.{χ->0};
Δ1=Δ/.{χ->z};
```

we can compute

```
TUa=TLa-Log[Δa/(1+θ(ζ-1))];
TU=TL-Log[Δ1/(1+θ(ζ-1))];
```

In a similar vein, we can compute T_S, relying on the calibrated values of T_M:

```
TS1=(aM*TM+aM*Log[vM]+aF*Log[vF]-(aM+aF)*Log[vS])/aS;
TSa=TMa;
TS=TS1/.{χ->z};
```

In a final step, we determine the income and unemployment effects of offshoring by computing

```
nΦ=Φ/.{χ->z};
nu=ua*Λ/.{χ->z};
```

To complete the calibration exercise, we finally add

```
Print["χ=", z, " AM= ", 100 (nAM-1), " AL= ", 100 (nAL-1)," ΔTS=", 100*(TS -
TSa)/TSa, " ΔTU=", 100*(TU-TUa)/TUa, " ΔT=", 100*(T-Ta)/Ta," I= ", 100*(nΦ-1),
" u= ", 100*(nu-ua)];
,{z0, {0.001, 0.01,0.1, 0.25, 0.5, 0.75, 0.9, N[χG]}}] ■
```

A.17 Proof of Proposition 4.1.1

Note that $\gamma = \Omega_S^{\alpha}\Omega_N^{1-\alpha} < 1$ in Eq. (4.1′) still endogenously depends on domestic factor prices, w_S and w_N, respectively. In order to obtain a testable prediction on how falling offshoring costs, τ_S and τ_N, relate to the individual skill upgrading decision of domestic workers in Eq. (4.1′) we have to replace w_S and w_N in Ω_S and Ω_N. Using the definitions of $\Omega_S = (\tau_S w_S^*/w_S)^{1-\theta} \leq 1$ and $\Omega_N = (\tau_N w_N^*/w_N)^{1-\theta} \leq 1$, we replace w_S and w_N by Eqs. (4.2) and (4.3). Skill upgrading condition (4.1′) can then be written as

$$u = \frac{\alpha s^{\alpha-1} - (1-\alpha)\, s^{\alpha}}{\left[A\,(\tau_S w_S^*)^{\alpha}\,(\tau_N w_N^*)^{1-\alpha}\,\Omega_S^{\alpha}\Omega_N^{1-\alpha}\right]^{(1-\theta)}} - \kappa,$$

in which $A \equiv 1/[\alpha^{\alpha}\,(1-\alpha)^{1-\alpha}] > 0$ is a positive constant. Unfortunately the above expression still depends on $\gamma = \Omega_S^{\alpha}\Omega_N^{1-\alpha}$. However, replacing again w_S and w_N in Ω_S and Ω_N by Eqs. (4.2) and (4.3), we find that after K iterations Eq. (4.1′) can be rewritten as a sequence $Z\,(K)$ with

$$u \equiv Z\,(K) = \frac{\alpha s^{\alpha-1} - (1-\alpha)\, s^{\alpha}}{\left[A\,(\tau_S w_S^*)^{\alpha}\,(\tau_N w_N^*)^{1-\alpha}\right]^{\sum_{k=1}^{K}(1-\theta)^k}\left(\Omega_S^{\alpha}\Omega_N^{1-\alpha}\right)^{(1-\theta)^K}} - \kappa.$$

Letting K go to infinity we find that sequence $Z(K)$ converges to

$$\lim_{K\to\infty} Z(K) = \frac{\alpha s^{\alpha-1} - (1-\alpha)s^{\alpha}}{\left[A(\tau_S w_S^*)^{\alpha}(\tau_N w_N^*)^{1-\alpha}\right]^{\frac{1-\theta}{\theta}}} - \kappa,$$

as $\lim_{K\to\infty}\sum_{k=1}^{K}(1-\theta)^k = (1-\theta)/\theta$. The above equation no longer depends on w_S and w_N such that it is easy to infer that $\partial s/\partial\tau_S < 0$ and $\partial s/\partial\tau_N < 0$. Thus, a gradual decline in *any* offshoring cost, τ_S or τ_N, leads to more individual skill upgrading. Taking additionally into account that according to Eq. (4.4) the share of tasks performed domestically is proportional to the respective cost savings factor from offshoring, $\Omega_S \leq 1$ or $\Omega_N \leq 1$, proposition 4.1.1 follows immediately. ■

A.18 Proof of Lemma 5.2.1

In order to prove Lemma 5.2.1 it suffices to show that given the production function in Eq. (5.1) firms optimally decide to match only workers of the same expected skill, such that $\bar{s}_l = \bar{s}_\ell$ with $l = 1,2$ and $\ell \in \{L, H\}$. The simple proof presented here is taken from Basu (1997). For a more general proof of positive assortative matching see Becker (1991) or Sattinger (1975).

Consider two different arbitrary average skill levels, $\bar{s}_L$ and $\bar{s}_H$, with $\bar{s}_H > \bar{s}_L$. A firm facing the optimisation problem in Eq. (5.2) now has three different possibilities of pairing workers:

$$\pi(\bar{s}_H, \bar{s}_H) = 2A\bar{s}_H^2 - 2w(\bar{s}_H), \tag{A.76}$$

$$\pi(\bar{s}_L, \bar{s}_L) = 2A\bar{s}_L^2 - 2w(\bar{s}_L), \tag{A.77}$$

$$\pi(\bar{s}_H, \bar{s}_L) = 2A\bar{s}_H\bar{s}_L - w(\bar{s}_H) - w(\bar{s}_L). \tag{A.78}$$

Let us first suppose $\pi(\bar{s}_H, \bar{s}_L) \geq \pi(\bar{s}_H, \bar{s}_H)$ which results in the following chain of inequalities

$$2A\bar{s}_H\bar{s}_L - w(\bar{s}_H) - w(\bar{s}_L) \geq 2A\bar{s}_H^2 - 2w(\bar{s}_H), \tag{A.79}$$

$$2A\bar{s}_H(\bar{s}_H - \bar{s}_L) \leq w(\bar{s}_H) - w(\bar{s}_L), \tag{A.80}$$

$$2A\bar{s}_L(\bar{s}_H - \bar{s}_L) < w(\bar{s}_H) - w(\bar{s}_L), \tag{A.81}$$

$$2A\bar{s}_H\bar{s}_L - w(\bar{s}_H) - w(\bar{s}_L) < 2A\bar{s}_L^2 - 2w(\bar{s}_L), \tag{A.82}$$

where $\bar{s}_H > \bar{s}_L$ has been utilised to derive the inequality in Eq. (A.81) from Eq. (A.80). Note that the inequality in Eq. (A.82) implies $\pi(\bar{s}_L, \bar{s}_L) \geq \pi(\bar{s}_H, \bar{s}_L)$. Now imagine $\pi(\bar{s}_L, \bar{s}_H) \geq \pi(\bar{s}_L, \bar{s}_L)$

giving rise to

$$2A\bar{s}_H\bar{s}_L - w(\bar{s}_H) - w(\bar{s}_L) \geq 2A(\bar{s}_L)^2 - 2w(\bar{s}_L), \tag{A.83}$$

$$2A\bar{s}_L(\bar{s}_H - \bar{s}_L) \geq w(\bar{s}_H) - w(\bar{s}_L), \tag{A.84}$$

$$2A\bar{s}_H(\bar{s}_H - \bar{s}_L) > w(\bar{s}_H) - w(\bar{s}_L), \tag{A.85}$$

$$2A(\bar{s}_H)^2 - 2w(\bar{s}_H) > 2A\bar{s}_H\bar{s}_L - w(\bar{s}_H) - w(\bar{s}_L), \tag{A.86}$$

where again $\bar{s}_H > \bar{s}_L$ has been utilised to derive the inequality in Eq. (A.85) from Eq. (A.84). The inequality in Eq. (A.86) implies $\pi(\bar{s}_H, \bar{s}_H) \geq \pi(\bar{s}_H, \bar{s}_L)$. Taking stock, profits from positive assortative matching always surpass profits from cross matching, such that firms always decide to pair workers of identical skill, i.e. $\bar{s}_l = \bar{s}_\ell$. ■

A.19 Proof of Proposition 5.2.2

The proof proceeds in two steps. At first we prove the existence and uniqueness of an equilibrium with positive selection into migration, i.e. $\tilde{s} \in (0,1)$ for $c \in (0, A/2)$. Then it is shown that the equilibrium with no migration is unique provided that $c \geq A/2$.

For the equilibrium with $c \in (0, A/2)$ to exist, it must hold that at least one worker has an incentive to deviate, when being in an equilibrium without migration, while no worker wants to deviate, when being in an equilibrium with migration. More formally, focussing on the worker with the maximum skill, $s = 1$, it can be shown that $w(\bar{s}, 1) = A/2 < w(1,1) - c = A - c \;\; \forall \;\; c \in (0, A/2)$. Conversely, it follows for all workers, $s \in [\tilde{s}, 1]$, that $w(\bar{s}_L, s) \leq w(\bar{s}_H, s) - c \;\; \forall \;\; c = \tilde{s}A/2 \in (0, A/2)$. To see this, note that the above inequality simplifies to $\tilde{s} \leq s$, once $c = \tilde{s}A/2$ as well as $\bar{s}_L = \tilde{s}/2$ and $\bar{s}_L = (1 + \tilde{s})/2$ are replaced. Uniqueness can be proven by considering a simple counter example which readily extents to more general cases. Imagine two (or more) cutoffs, $0 < \tilde{s}_1 < \tilde{s}_2 < 1$, exist. Moreover assume all workers with $s \in [0, \tilde{s}_1) \;\; \wedge \;\; s \in [\tilde{s}_2, 1]$ and average skill, $\bar{s}_1$, stay in Home, while all workers with $s \in [\tilde{s}_1, \tilde{s}_2)$ and average skill, $\bar{s}_2$, select into migration. For the workers with critical abilities, $\tilde{s}_1$ and $\tilde{s}_2$, then must hold that $w(\bar{s}_1, \tilde{s}_1) = w(\bar{s}_2, \tilde{s}_1) - c$ and $w(\bar{s}_1, \tilde{s}_2,) = w(\bar{s}_2, \tilde{s}_2) - c$, respectively. Since $w(\bar{s}_\ell, s) \;\; \forall \;\; \ell = 1, 2$ increases monotonically in s, the above two conditions can only be fulfilled simultaneously, if $\tilde{s}_1 = \tilde{s}_2 = \tilde{s}$, which rules out multiplicity of equilibria.

The equilibrium without migration is stable for $c \geq A/2$, if no individual with skill, $s \in [0,1]$, has an incentive to deviate, i.e. even if deviation allows to match with the most high-skilled co-worker with $s = 1$, the condition, $w(\bar{s}, s) \geq w(1, s) - c$, should hold. Replacing $c \geq A/2$ this condition simplifies to $1 \geq s$, which is fulfilled for all $s \in [0,1]$. ■

A.20 Individual welfare effects and the distribution of skills

Instead of an uniform distribution with support $s \in [0,1]$ we assume an arbitrary distribution with the same support and cumulative density function $F(s)$. We furthermore assume $F(0) = 0$ as well as $F(1) = 1$ and denote the corresponding density function by $f(s) = \partial F(s)/\partial s$. The average skills within the group of low-skilled natives and high-skilled migrants then can be expressed as

$$\bar{s}_L = \frac{\int_0^{\tilde{s}} s f(s)\, ds}{F(\tilde{s})}, \qquad \text{and} \qquad \bar{s}_H = \frac{\int_{\tilde{s}}^1 s f(s)\, ds}{1 - F(\tilde{s})}, \tag{A.87}$$

where $\tilde{s}$, as before, denotes the critical worker, separating both groups. The migration indifference condition remains unchanged and is given by $A\bar{s}_L\tilde{s}^{\mathrm{lf}} = A\bar{s}_H\tilde{s}^{\mathrm{lf}} - c$. Focussing on the expected net wage of the most high-skilled workers with $s = 1$, we obtain

$$w_H(\tilde{s}^{\mathrm{lf}}, 1) - c = A\frac{\int_{\tilde{s}^{\mathrm{lf}}}^1 s f(s)\, ds}{1 - F(\tilde{s}^{\mathrm{lf}})} - c, \tag{A.88}$$

$$= A\frac{\int_{\tilde{s}^{\mathrm{lf}}}^1 s f(s)\, ds}{1 - F(\tilde{s}^{\mathrm{lf}})}(1 - \tilde{s}^{\mathrm{lf}}) + A\frac{\int_0^{\tilde{s}^{\mathrm{lf}}} s f(s)\, ds}{F(\tilde{s}^{\mathrm{lf}})}\tilde{s}^{\mathrm{lf}}, \tag{A.89}$$

which is larger than the expected autarky wage $A\int_0^1 s f(s)\, ds$ if and only if $[F(\tilde{s}^{\mathrm{lf}}) - \tilde{s}^{\mathrm{lf}}][\bar{s}_H - \bar{s}_L] > 0$. Since $\bar{s}_H \geq \bar{s}_L$, we have $w_H(\tilde{s}^{\mathrm{lf}}, 1) - c < A\int_0^1 s f(s)\, ds$ if $F(\tilde{s}^{\mathrm{lf}}) < \tilde{s}^{\mathrm{lf}}$ and $w_H(\tilde{s}^{\mathrm{lf}}, 1) - c > A\int_0^1 s f(s)\, ds$ if $F(\tilde{s}^{\mathrm{lf}}) > \tilde{s}^{\mathrm{lf}}$. ■

A.21 Redistribution of the emigration tax revenue

Let the social planner equilibrium be implemented by (symmetric) emigration taxes $t^{\mathrm{sp}} = A/4$. Governments only care about their nationals and redistribute all tax revenue equally to them irrespective of their residence. We show that even with redistribution the marginal migrant $\tilde{s}^{\mathrm{sp}}$ loses relative to autarky. To see this, note that for the marginal migrant to benefit from migration and redistribution

$$w\left(1, \tilde{s}^{\text{sp}}\right) \leq w_H\left(\tilde{s}^{\text{sp}}, \tilde{s}^{\text{sp}}\right) - c - t^{\text{sp}} + \left(1 - \tilde{s}^{\text{sp}}\right) t^{\text{sp}}, \tag{A.90}$$

has to hold, where $w_H\left(\tilde{s}^{\text{sp}}, \tilde{s}^{\text{sp}}\right) - c - t^{\text{sp}}$ is the marginal migrant's net wage after having paid cost c and tax t^{sp}, while $w\left(1, \tilde{s}^{\text{sp}}\right)$ is the autarky wage paid to an individual with comparable skill $\tilde{s}^{\text{sp}}$. As population size is normalized to unity, total tax revenue $\left(1 - \tilde{s}^{\text{sp}}\right) t^{\text{sp}}$ equals the amount of redistribution each individual receives. Using $w_\ell(\bar{s}_\ell, s) = A\bar{s}_\ell s \ \forall \ \ell = L, H$ in combination with $\bar{s}_L = \tilde{s}^{\text{sp}}/2$, $\bar{s}_H = (1 + \tilde{s}^{\text{sp}})/2$ as well as $\tilde{s}^{\text{sp}} = 2\left(c + t^{\text{sp}}\right)/A$ and $t^{\text{sp}} = A/4$ allows us to rewrite Eq. (A.90) as:

$$\frac{c}{4} \leq \frac{c^2}{A}. \tag{A.91}$$

Eq. (A.91) holds with equality for $c = 0$, and collapses otherwise to $c > A/4$, which contradicts $\tilde{s}^{\text{sp}} = \frac{1}{2} + \frac{2c}{A} < 1$.

Let us now extend this result to individuals in the neighbourhood of $\tilde{s}^{\text{sp}}$. We focus on workers $0 \leq s_1^{\text{sp}} \leq \tilde{s}^{\text{sp}} \leq \tilde{s}_2^{\text{sp}} \leq 1$ who are indifferent between the outcomes in the social planner equilibrium with redistribution and autarky, such that:

$$w\left(1, s_1^{\text{sp}}\right) = w_L\left(\tilde{s}^{\text{sp}}, s_1^{\text{sp}}\right) + t^{\text{sp}}\left(1 - \tilde{s}^{\text{sp}}\right) \tag{A.92}$$

$$w\left(1, s_2^{\text{sp}}\right) = w_H\left(\tilde{s}^{\text{sp}}, s_2^{\text{sp}}\right) - c - t^{\text{sp}} + t^{\text{sp}}\left(1 - \tilde{s}^{\text{sp}}\right) \tag{A.93}$$

Again using $w_\ell(\bar{s}_\ell, s) = A\bar{s}_\ell s \ \forall \ \ell = L, H$ in combination with $\bar{s}_L = \tilde{s}^{\text{sp}}/2$, $\bar{s}_H = (1 + \tilde{s}^{\text{sp}})/2$ as well as $\tilde{s}^{\text{sp}} = 2\left(c + t^{\text{sp}}\right)/A$ and $t^{\text{sp}} = A/4$ allows us to solve Eqs. (A.92) and (A.93) for:

$$s_1^{\text{sp}} = \frac{1}{2} \quad \text{and} \quad s_2^{\text{sp}} = \frac{\frac{1}{2} + \frac{6c}{A}}{1 + \frac{4c}{A}} \in [1/2, 1] \ \ \forall \ c \in [0, A/4]\,. \tag{A.94}$$

Note that $s_1^{\text{sp}} < \tilde{s}^{\text{sp}} < s_2^{\text{sp}}$, if $c \in (0, A/4)$. Thus, after redistribution the most able non-migrants as well as the least skilled migrants are worse off than under autarky. ■

A.22 Proof of Proposition 5.4.4

Without loss of generality let us focus on $c_e > c_m$, the opposite case, $c_m > c_e$, then follows by symmetry. Equilibrium may feature up to three groups of workers: migrants, educated workers, and those who do not signal at all. The corresponding subindices are $\hat{\ell} \in \{m, e, n\}$. As in the baseline model the wage rate of a worker with skill $s_{\hat{\ell}}$ is given by Eq. (5.3), i.e. $w\left(\bar{s}_{\hat{\ell}}, s_{\hat{\ell}}\right) = A\bar{s}_{\hat{\ell}}s_{\hat{\ell}}$ with $\bar{s}_{\hat{\ell}}$ being the average skill in group $\hat{\ell} \in \{m, e, n\}$. At prohibitive cost, $c_m \geq A/2$ and /or $c_e \geq A/2$, no worker with $s \in [0, 1]$ has an incentive to deviate from an equilibrium without migration. To see this, note that $w(\bar{s}, s) \geq w(s, s) - c_m$, in which $\bar{s} = 1/2$, can be rewritten as $1 \geq s \ \forall \ s \in [0, 1]$, once we take into account that $c_m \geq A/2$. An analogous logic applies to education, if $c_e \geq A/2$.

For scenario *(i)* to arise the sorting $0 \leq \tilde{s}_m < \tilde{s}_e \leq 1$ has to apply. Note that all workers with $s \in [0, \tilde{s}_m)$ choose not to signal, all workers with $s \in [\tilde{s}_m, \tilde{s}_e)$ select into migration and all workers with $s \in [\tilde{s}_e, 1]$ select into education. The resulting average skills in the three groups are then given by $\bar{s}_n = \tilde{s}_m/2$, $\bar{s}_m = (\tilde{s}_m + \tilde{s}_e)/2$ and $\bar{s}_e = (\tilde{s}_e + 1)/2$. Intuitively, $\bar{s}_e > \bar{s}_m > \bar{s}_n$. For the marginal workers, $\tilde{s}_m$ and $\tilde{s}_e$ the following two arbitrage conditions apply

$$w(\bar{s}_n, \tilde{s}_m) = w(\bar{s}_m, \tilde{s}_m) - c_m, \tag{A.95}$$

$$w(\bar{s}_m, \tilde{s}_e) - c_m = w(\bar{s}_e, \tilde{s}_e) - c_e. \tag{A.96}$$

Solving both equations yields $\tilde{s}_e = 2c_e/A$ and $\tilde{s}_m = c_m/c_e$ for all $c_m < 2c_e^2/A < A/2$. Note that the first inequality sign in this condition follows from $\tilde{s}_m < \tilde{s}_e$, while the second inequality sign follows from $\tilde{s}_e < 1$.

Scenario *(ii)* is derived in a stepwise fashion by eliminating all possible equilibrium sortings, except for $0 \leq \tilde{s}_e \leq \tilde{s}_m = 1$. Intuitively, $0 \leq \tilde{s}_m < \tilde{s}_e \leq 1$ cannot be an equilibrium, since for $2c_e^2/A < c_m < c_e < A/2$, this sorting is incompatible with the incentive constraints in Eqs. (A.95) and (A.96). Assuming $0 \leq \tilde{s}_e < \tilde{s}_m \leq 1$ would imply that $\bar{s}_m > \bar{s}_e$. Recalling that $c_e > c_m$, this cannot be a stable outcome, as workers with $s \in [\tilde{s}_e, \tilde{s}_m)$, who select into education, bear larger cost, $c_e \geq c_m$, and have smaller gains, $w(s, \bar{s}_m) > w(s, \bar{s}_e)$, than those who select into migration. Clearly, for any worker with $s \in [\tilde{s}_e, \tilde{s}_m)$ using migration instead of education as a signalling device would be optimal. Since we already have shown that the equilibrium without migration does not exist for $c_m < c_e < A/2$, the only remaining equilibrium constellations are $0 \leq \tilde{s}_m \leq \tilde{s}_e = 1$ and $0 \leq \tilde{s}_e \leq \tilde{s}_m = 1$. Intuitively, individual incentives are such that each worker who is inclined to signal, prefers to be in the same group as the workers with the maximum skill, $s = 1$. On the contrary, the workers with $s = 1$ are indifferent between both signals, since in both equilibria welfare equals $w(\bar{s}_m, 1) - c_m = w(\bar{s}_e, 1) - c_e = A/2$. Extending this comparison to all workers with $s \in [\tilde{s}_e, 1]$, it can be easily shown that $w(\bar{s}_m, s) - c_m \leq w(\bar{s}_e, s) - c_e$. Thus, all workers with $s \in [\tilde{s}_e, 1]$ weakly prefer $0 \leq \tilde{s}_e \leq \tilde{s}_m = 1$ over $0 \leq \tilde{s}_m \leq \tilde{s}_e = 1$, such that $0 \leq \tilde{s}_e \leq \tilde{s}_m = 1$ is the only equilibrium for $2c_e^2/A < c_m < c_e < A/2$. ■

A.23 Derivation of Eqs. (5.11) and (5.12)

Facing the production function from Eq. (5.10), and knowing only the average skill within the two groups, L and H, firms optimally match together only workers from the same subgroup, L or H. We can therefore separately write down the reduced form profit maximization problem for the resulting two types of firms:

$$\max_{k_\ell} \; \pi_\ell(k_\ell) = 2A\bar{s}_\ell^2 k_\ell^\alpha - 2w(\bar{s}_\ell) - rk_\ell \;\; \forall \;\; \ell \in \{L, H\}. \tag{A.97}$$

The profit maximising level of capital depends on whether the firm employs individuals from group H or L. It follows from

$$\frac{\partial \pi_\ell(k_\ell)}{\partial k_\ell} = 2A\alpha\bar{s}_\ell^2 k_\ell^{\alpha-1} - r \stackrel{!}{=} 0 \;\; \forall \;\; \ell \in \{L, H\}, \tag{A.98}$$

and we get the standard result that the rate of return to capital, r, equals its value marginal product. Using the above equation in combination with the full employment condition:

$$\bar{k} = \tilde{s}k_L + (1-\tilde{s})k_H, \tag{A.99}$$

as well as $\bar{s}_L = \tilde{s}/2$ and $\bar{s}_H = (1+\tilde{s})/2$ allows us to solve for the amount of capital used by firms solely employing natives or migrants, respectively:

$$k_L = \left[\tilde{s} + (1-\tilde{s})\left(\frac{1+\tilde{s}}{\tilde{s}}\right)^{\frac{2}{1-\alpha}}\right]^{-1} \bar{k}, \tag{A.100}$$

$$k_H = \left[(1-\tilde{s}) + \tilde{s}\left(\frac{\tilde{s}}{1+\tilde{s}}\right)^{\frac{2}{1-\alpha}}\right]^{-1} \bar{k}, \tag{A.101}$$

with $k_H \geq \bar{k} \geq k_L$. ■

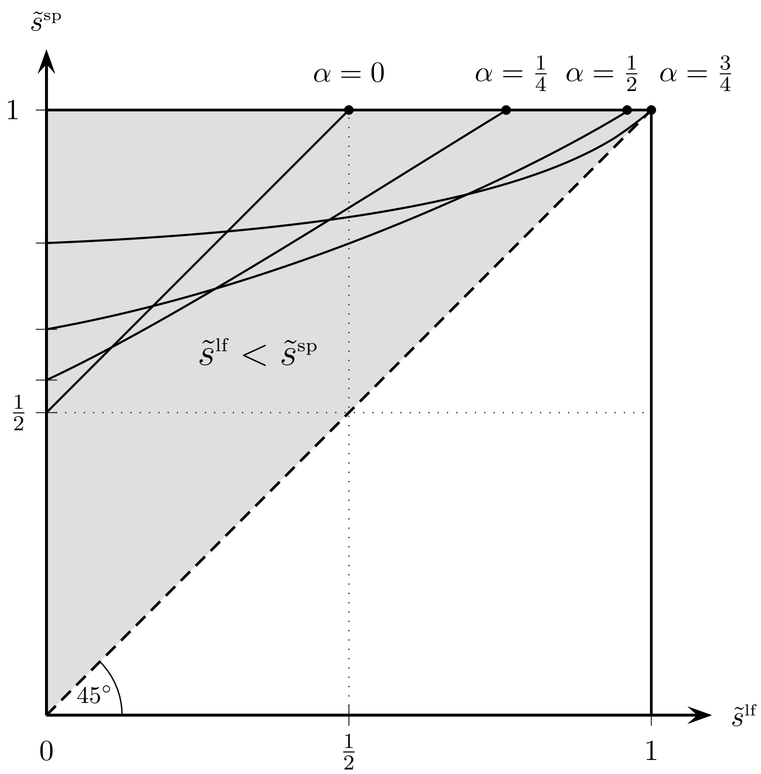
A.24 The social planner solution in a model with capital

We show that $\tilde{s}^{\mathrm{sp}} \geq \tilde{s}^{\mathrm{lf}}$ for all $\alpha \in [0, 1)$. Note that $\tilde{s}^{\mathrm{lf}}$ is implicitly determined by $w(\bar{s}_H, \tilde{s}^{\mathrm{lf}}) - w(\bar{s}_L, \tilde{s}^{\mathrm{lf}}) = c = A(1-\alpha)\tilde{s}^{\mathrm{lf}}(\bar{s}_H k_H^\alpha - \bar{s}_L k_L^\alpha)$. Substituting $\bar{s}_L = \tilde{s}^{\mathrm{lf}}/2$ and $\bar{s}_H = (1+\tilde{s}^{\mathrm{lf}})/2$, and replacing k_L and k_H by Eqs. (5.11) and (5.12) this indifference condition may be expressed as $\Omega(\tilde{s}^{\mathrm{lf}};\alpha) = c/A\bar{k}^{\alpha}$ where

$$\begin{aligned}\Omega(\tilde{s}^{\mathrm{lf}};\alpha) \equiv{}& (1-\alpha)\left[(1-\tilde{s}^{\mathrm{lf}}) + \tilde{s}^{\mathrm{lf}}\left(\frac{\tilde{s}^{\mathrm{lf}}}{1+\tilde{s}^{\mathrm{lf}}}\right)^{\frac{2}{1-\alpha}}\right]^{-\alpha}\frac{\tilde{s}^{\mathrm{lf}}\left(1+\tilde{s}^{\mathrm{lf}}\right)}{2} \\ &-(1-\alpha)\left[\tilde{s}^{\mathrm{lf}} + (1-\tilde{s}^{\mathrm{lf}})\left(\frac{1+\tilde{s}^{\mathrm{lf}}}{\tilde{s}^{\mathrm{lf}}}\right)^{\frac{2}{1-\alpha}}\right]^{-\alpha}\frac{(\tilde{s}^{\mathrm{lf}})^2}{2}.\end{aligned} \tag{A.102}$$

In order to obtain an implicit definition of $\tilde{s}^{\mathrm{sp}}$ we start out from aggregate welfare which is

Figure A.1: *Migration in the laissez faire and the social planner equilibrium*



given by $W(\tilde{s}) = \int_0^{\tilde{s}} A\bar{s}_L k_L^{\alpha} ds + \int_{\tilde{s}}^1 A\bar{s}_H k_H^{\alpha} ds - \int_{\tilde{s}}^1 c\ ds$. Again, substituting $\bar{s}_L = \tilde{s}^{\mathrm{lf}}/2$ and

$\bar{s}_H = (1+\tilde{s}^{\text{lf}})/2$, and replacing k_L and k_H by Eqs. (5.11) and (5.12) we obtain

$$\begin{aligned} W(\tilde{s}) = \frac{A}{4}\bar{k}^{\alpha} \Bigg\{ & \tilde{s}^3 \left[\tilde{s} + (1-\tilde{s}) \left(\frac{1+\tilde{s}}{\tilde{s}} \right)^{\frac{2}{1-\alpha}} \right]^{-\alpha} \\ & + (1-\tilde{s})(1+\tilde{s})^2 \left[(1-\tilde{s}) + \tilde{s} \left(\frac{\tilde{s}}{1+\tilde{s}} \right)^{\frac{2}{1-\alpha}} \right]^{-\alpha} \Bigg\} - (1-\tilde{s})\, c. \end{aligned} \tag{A.103}$$

In a next step we compute $dW(\tilde{s}^{\text{sp}})/d\tilde{s}^{\text{sp}} \overset{!}{=} 0$ and solve for $\Lambda(\tilde{s}^{\text{sp}};\alpha) = c/A\bar{k}^{\alpha}$ where

$$\Lambda(\tilde{s}^{\text{sp}};\alpha) \equiv - \left[\frac{dW(\tilde{s}^{\text{sp}})}{d\tilde{s}^{\text{sp}}} - c \right] \frac{1}{A\bar{k}^{\alpha}}. \tag{A.104}$$

Setting $\Omega(\tilde{s}^{\text{lf}};\alpha) = \Lambda(\tilde{s}^{\text{sp}};\alpha)$ we finally obtain a link between $\tilde{s}^{\text{lf}}$ and $\tilde{s}^{\text{sp}}$ for the same economic fundamentals A, $\bar{k}$, and c. Plotting this link for varying levels of $\alpha \in [0,1)$ in Figure A.1 reveals that $\tilde{s}^{\text{sp}} \geq \tilde{s}^{\text{lf}}$ for all $\alpha \in [0,1)$. In particular we show that for $\alpha = 0$ we get the familiar result $\tilde{s}^{\text{sp}} = \tilde{s}^{\text{lf}} + \frac{1}{2}$. ■

A.25 Proof of Proposition 5.4.6

We start by proving that $W_i(\tilde{s}_i^{\text{lf}}) \leq W_i(1) = A_i/4 \;\forall\; A_i \geq A_j$. Note that for $\tilde{s}_i^{\text{lf}} = 1$ we obtain $W_i(1) = A_i/4$ from Eq. (5.16). Using this together with $W_i(\tilde{s}_i^{\text{lf}})$ from Eq. (5.16) we obtain

$$\begin{aligned} W_i(\tilde{s}_i^{\text{lf}}) - W_i(1) &= (A_j - A_i)\left[1 + (\tilde{s}_i^{\text{lf}})^3\right] - A_i(1+\tilde{s}_i^{\text{lf}})\,\tilde{s}_i^{\text{lf}} + A_i 2\,(\tilde{s}_i^{\text{lf}})^2 \\ &\leq (A_j - A_i)\left[1 + (\tilde{s}_i^{\text{lf}})^3 - 2\,(\tilde{s}_i^{\text{lf}})^2\right], \end{aligned} \tag{A.105}$$

where the last line is non-positive whenever $1 + (\tilde{s}_i^{\text{lf}})^3 - 2\,(\tilde{s}_i^{\text{lf}})^2 \geq 0$ and $A_i \geq A_j$. Since $1 + (\tilde{s}_i^{\text{lf}})^3 - 2(\tilde{s}_i^{\text{lf}})^2$ has a local maximum at $\tilde{s}_i^{\text{lf}} = 0$ and intersects the abscissa at $\tilde{s}_i^{\text{lf}} = 1$ and $\tilde{s}_i^{\text{lf}} = 1/2 \pm \sqrt{5/4}$, we have $1 + (\tilde{s}_i^{\text{lf}})^3 - 2(\tilde{s}_i^{\text{lf}})^2 \geq 0 \;\forall\; \tilde{s}_i^{\text{lf}} \in [0,1]$ and, hence, $W_i(\tilde{s}_i^{\text{lf}}) \leq W_i(1) \;\forall\; A_i \geq A_j$. In order to complete the proof of Proposition 5.4.6 it remains to show that for $A_j > A_i$ we have $W_i^{\text{lf}}(c) \geq A_i/4 \;\forall\; 0 \leq c \leq \frac{1}{4}\left(2A_j - 3A_i + \sqrt{4A_j^2 - 8A_iA_j + 5A_i^2}\right)$, while $W_i^{\text{lf}}(c) < A_i/4 \;\forall\; \frac{1}{4}\left(2A_j - 3A_i + \sqrt{4A_j^2 - 8A_iA_j + 5A_i^2}\right) < c < \frac{1}{2}(2A_j - A_i)$. Using Eq. (5.15) to substitute for $\tilde{s}_i^{\text{lf}}$ in Eq. (5.16), it can be shown that $W_i^{\text{lf}}(c) - A_i/4 = 0$ has three solutions, which are

$$c_1 = \frac{1}{2}\left(2A_j - A_i\right), \tag{A.106}$$

$$c_2 = \frac{1}{4}\left(2A_j - 3A_i + \sqrt{4A_j^2 - 8A_iA_j + 5A_i^2}\right), \tag{A.107}$$

$$c_3 = \frac{1}{4}\left(2A_j - 3A_i - \sqrt{4A_j^2 - 8A_iA_j + 5A_i^2}\right). \tag{A.108}$$

Note that since $2/3 < A_D/A_F < 3/2$, solution (A.108) is negative and therefore economically irrelevant. Solution (A.106) equals the prohibitive migration cost at which $\tilde{s}_i^{\text{lf}} = 1$. Finally, it is easily checked that $0 < \frac{1}{4}\left(2A_j - 3A_i + \sqrt{4A_j^2 - 8A_iA_j + 5A_i^2}\right) < \frac{1}{2}\left(2A_j - A_i\right)$. Since Eq. (5.16) implies $W_i^{\text{lf}}(0) = A_j/4 > A_i/4$ we can immediately infer that for low migration cost, i.e. $0 \leq c \leq \frac{1}{4}\left(2A_j - 3A_i + \sqrt{4A_j^2 - 8A_iA_j + 5A_i^2}\right)$ aggregate welfare gains exist, while for high migration cost, $\frac{1}{4}\left(2A_j - 3A_i + \sqrt{4A_j^2 - 8A_iA_j + 5A_i^2}\right) < c < \frac{1}{2}\left(2A_j - A_i\right)$, aggregate losses result. ■

A.26 The social planner solution in a model with asymmetric countries

In order to show that $\tilde{s}_i^{\text{sp}} > \tilde{s}_i^{\text{lf}}\ \forall\ i, j \in \{D, F\}$ with $i \neq j$ we have to consider two scenarios in which either $A_i > A_j$ or $A_i < A_j$. We start with $A_i < A_j$ and rewrite $\tilde{s}_i^{\text{sp}} > \tilde{s}_i^{\text{lf}}$ as

$$-4c\left(A_j - c\right)\left(A_j - A_i\right) < A_iA_j^2. \tag{A.109}$$

Note that Eq. (A.109) holds true for the economically relevant parameter space $0 \leq c \leq \frac{1}{2}\left(2A_j - A_i\right)$, which completes the first part of the proof. Now suppose $A_i > A_j$, such that $\tilde{s}_i^{\text{sp}} > \tilde{s}_i^{\text{lf}}$ can be expressed as

$$-4c\left(A_j - c\right)\left(A_j - A_i\right) - A_iA_j^2 < -\frac{4}{3}c\left(A_j - c\right) - A_j^2 < \frac{4}{3}c^2 - A_j^2 < 0, \tag{A.110}$$

where we have made use of the fact that $\frac{2}{3}A_i < A_j < A_i$. The last inequality in (A.110) holds true, since the economically relevant parameter space $0 \leq c \leq \frac{1}{2}\left(2A_j - A_i\right)$ translates into $0 \leq c \leq \frac{1}{2}A_j$ for $\frac{2}{3}A_i < A_j < A_i$, such that we can be sure that for $0 \leq c \leq \frac{1}{2}\left(2A_j - A_i\right)$ we have $\tilde{s}_i^{\text{sp}} > \tilde{s}_i^{\text{lf}}$. ■

A.27 Proof of Lemma 5.6.1

In order to prove Lemma 5.6.1 it suffices to show that given our production function in Eq. (5.18) firms optimally decide to match only workers of the same expected skill, such that $\bar{s}_l = \bar{s}_{\ell t}$ with $l = 1, 2$; $\ell \in \{N, M, R\}$ and $t = 1, 2$. The simple proof presented here is based on Basu (1997) and extends our proof from Appendix A.18 to a scenario with arbitrary many subgroups $(\ell \times t)$. For a more general proof of positive assortative matching see Becker (1991) or Sattinger (1975).

Consider two arbitrary chosen average skill levels $\bar{s}_{\ell t} \neq \bar{s}_{\ell' t'}$ with $\ell, \ell' \in \{N, M, R\}$ and $t, t' = 1, 2$ as well as $\ell \neq \ell'$ and $t \neq t'$. Now let us – without loss of generality – assume that $\bar{s}_{\ell t} > \bar{s}_{\ell' t'}$. A firm with two vacancies has three different possibilities of pairing the workers from the subgroups $\ell \times t$ and $\ell' \times t'$ in order to maximise Eq. (5.19):

$$\pi\left(\bar{s}_{\ell t}, \bar{s}_{\ell t}\right) = 2A\bar{s}_{\ell t}^2 - 2w\left(\bar{s}_{\ell t}\right), \tag{A.111}$$

$$\pi\left(\bar{s}_{\ell' t'}, \bar{s}_{\ell' t'}\right) = 2A\bar{s}_{\ell' t'}^2 - 2w\left(\bar{s}_{\ell' t'}\right), \tag{A.112}$$

$$\pi\left(\bar{s}_{\ell t}, \bar{s}_{\ell' t'}\right) = 2A\bar{s}_{\ell t}\bar{s}_{\ell' t'} - w\left(\bar{s}_{\ell t}\right) - w\left(\bar{s}_{\ell' t'}\right). \tag{A.113}$$

Let us first suppose $\pi\left(\bar{s}_{\ell t}, \bar{s}_{\ell' t'}\right) \geq \pi\left(\bar{s}_{\ell t}, \bar{s}_{\ell t}\right)$, which results in the following chain of inequalities:

$$2A\bar{s}_{\ell t}\bar{s}_{\ell' t'} - w\left(\bar{s}_{\ell t}\right) - w\left(\bar{s}_{\ell' t'}\right) \geq 2A\bar{s}_{\ell t}^2 - 2w\left(\bar{s}_{\ell t}\right), \tag{A.114}$$

$$2A\bar{s}_{\ell t}\left(\bar{s}_{\ell t} - \bar{s}_{\ell' t'}\right) \leq w\left(\bar{s}_{\ell t}\right) - w\left(\bar{s}_{\ell' t'}\right), \tag{A.115}$$

$$2A\bar{s}_{\ell' t'}\left(\bar{s}_{\ell t} - \bar{s}_{\ell' t'}\right) < w\left(\bar{s}_{\ell t}\right) - w\left(\bar{s}_{\ell' t'}\right), \tag{A.116}$$

$$2A\bar{s}_{\ell t}\bar{s}_{\ell' t'} - w\left(\bar{s}_{\ell t}\right) - w\left(\bar{s}_{\ell' t'}\right) < 2A\bar{s}_{\ell' t'}^2 - 2w\left(\bar{s}_{\ell' t'}\right), \tag{A.117}$$

where $\bar{s}_{\ell t} > \bar{s}_{\ell' t'}$ has been utilised to derive inequality (A.116) from (A.115). Note that inequality (A.117) then implies $\pi\left(\bar{s}_{\ell' t'}, \bar{s}_{\ell' t'}\right) \geq \pi\left(\bar{s}_{\ell t}, \bar{s}_{\ell' t'}\right)$. Now imagine $\pi\left(\bar{s}_{\ell' t'}, \bar{s}_{\ell t}\right) \geq \pi\left(\bar{s}_{\ell' t'}, \bar{s}_{\ell' t'}\right)$, giving rise to:

$$2A\bar{s}_{\ell t}\bar{s}_{\ell' t'} - w\left(\bar{s}_{\ell t}\right) - w\left(\bar{s}_{\ell' t'}\right) \geq 2A\bar{s}_{\ell' t'}^2 - 2w\left(\bar{s}_{\ell' t'}\right), \tag{A.118}$$

$$2A\bar{s}_{\ell' t'}\left(\bar{s}_{\ell t} - \bar{s}_{\ell' t'}\right) \geq w\left(\bar{s}_{\ell t}\right) - w\left(\bar{s}_{\ell' t'}\right), \tag{A.119}$$

$$2A\bar{s}_{\ell t}\left(\bar{s}_{\ell t} - \bar{s}_{\ell' t'}\right) > w\left(\bar{s}_{\ell t}\right) - w\left(\bar{s}_{\ell' t'}\right), \tag{A.120}$$

$$2A\left(\bar{s}_{\ell t}\right)^2 - 2w\left(\bar{s}_{\ell t}\right) > 2A\bar{s}_{\ell t}\bar{s}_{\ell' t'} - w\left(\bar{s}_{\ell t}\right) - w\left(\bar{s}_{\ell' t'}\right), \tag{A.121}$$

where again $\bar{s}_{\ell t} > \bar{s}_{\ell' t'}$ has been utilised to derive inequality (A.120) from (A.119). Inequality (A.121) now implies $\pi(\bar{s}_{\ell t}, \bar{s}_{\ell t}) \geq \pi(\bar{s}_{\ell t}, \bar{s}_{\ell' t'})$. Taking stock, profits from positive assortative matching always surpass profits from cross matching, such that firms always decide to pair workers of identical skill, i.e. $\bar{s}_l = \bar{s}_{\ell t}$. ■

A.28 Proof of Proposition 5.7.1

We show that in a migration equilibrium the lifetime income of *all* workers, i.e. wages $w(\bar{s}_{\ell t}, s)$ $\forall\ \ell \in \{N, M, R\}$ at age $t = 1, 2$ less the migration cost c, if applicable, is *not* higher than the comparable lifetime income $\sum_{t=1}^{2} w(\bar{s}_t, s) = 2w(\bar{s}, s)$ in a non-migration equilibrium, with $\bar{s}_t = \bar{s} = 1/2$ being the time-constant average skill of the overall population.

Beginning with the high cost $\hat{c} \in [\hat{c}_b^{\text{lf}}, \hat{c}_c^{\text{lf}})$ case (c) we demonstrate that

$$2w(\bar{s}, s) \geq \begin{cases} w(\bar{s}_{N1}, s) + w(\bar{s}_{N2}, s) & \text{if} \quad s < \tilde{s}_m^{\text{lf}}, \\ w(\bar{s}_{M1}, s) - c + w(\bar{s}_{R2}, s) & \text{if} \quad s \geq \tilde{s}_m^{\text{lf}}. \end{cases} \tag{A.122}$$

Using the definition of $w(\bar{s}_{\ell t}, s)$ in Eq. (5.20) in combination with $\bar{s} = 1/2 \geq \bar{s}_{Nt} = \tilde{s}_m^{\text{lf}}/2\ \forall\ t = 1, 2$ and $\tilde{s}_m^{\text{lf}} = \hat{c}$ from Eq. (5.26), we reduce the first inequality in Eq. (A.122) to $\hat{c} \leq 1$, which in case (c) with $\hat{c} < \hat{c}_c^{\text{lf}} = 1$ always is fulfilled. Substituting $\bar{s}_{M1} = \bar{s}_{R2} = (\tilde{s}_m^{\text{lf}} + 1)/2 \geq \bar{s} = 1/2$ and $\tilde{s}_m^{\text{lf}} = \hat{c}$ from Eq. (5.26) into the second inequality in Eq. (A.122) we can solve for $s \leq 1$, which generally holds true, since $s \in [0, 1]$.

Turning to the low cost $\hat{c} \in [\hat{c}_a^{\text{lf}}, \hat{b}_b^{\text{lf}})$ case (b) we show that

$$2w(\bar{s}, s) \geq \begin{cases} w(\bar{s}_{N1}, s) + w(\bar{s}_{N2}, s) & \text{if} \quad s < \tilde{s}_m^{\text{lf}}, \\ w(\bar{s}_{M1}, s) - c + w(\bar{s}_{R2}, s) & \text{if} \quad s \in [\tilde{s}_m^{\text{lf}}, \tilde{s}_r^{\text{lf}}), \\ w(\bar{s}_{M1}, s) - c + w(\bar{s}_{M2}, s) - c & \text{if} \quad s \geq \tilde{s}_r^{\text{lf}}. \end{cases} \tag{A.123}$$

Using $\bar{s} = 1/2 \geq \bar{s}_{Nt} = \tilde{s}_m^{\text{lf}}/2\ \forall\ t = 1, 2$ in combination with $\tilde{s}_m^{\text{lf}} = (1 + 4\hat{c} - \sqrt{1 + 16\hat{c}^2})/2$ from Eq. (5.26), we transform the first inequality in Eq. (A.123) into $\hat{c} \geq 0$, which in case (b) with $\hat{c} \geq \hat{c}_a^{\text{lf}} = 0$ always is fulfilled. Note that the second inequality in Eq. (A.123) can be rewritten as

$$\lambda(s) \equiv 1 - \tilde{s}_m^{\text{lf}} - \frac{1 + \tilde{s}_r^{\text{lf}}}{2} + \frac{\hat{c}}{s} \geq 0, \tag{A.124}$$

where $\bar{s} = 1/2 \leq \bar{s}_{M1} = (\tilde{s}_m^{\text{lf}} + 1)/2$ and $\bar{s}_{R2} = (\tilde{s}_m^{\text{lf}} + \tilde{s}_r^{\text{lf}})/2$ have been used to replace $\bar{s}$, $\bar{s}_{M1}$, and $\bar{s}_{R2}$. Since $\lambda'(s) < 0$, we have $\lambda(s) \geq \lambda(1)$ and $\lambda(1) \geq 0$ is sufficient for $\lambda(s) \geq 0$. Using $\tilde{s}_m^{\text{lf}} = (1 + 4\hat{c} - \sqrt{1 + 16\hat{c}^2})/2$ from Eq. (5.26) and $\tilde{s}_r^{\text{lf}} = 4\hat{c}/(1 - 4\hat{c} + \sqrt{1 + 16\hat{c}^2})$ from Eq. (5.27) we find that $\lambda(1) \geq 0$ may equivalently be expressed as $\hat{c} \leq 1/3$, which in case (b) with $\hat{c} < \hat{c}_b^{\text{lf}} = 1/3$ always is fulfilled. Replacing $\bar{s}$, $\bar{s}_{M1}$, and $\bar{s}_{M2}$ in the third inequality of Eq. (A.123) by $\bar{s} = 1/2 \leq \bar{s}_{M1} = (\tilde{s}_m^{\text{lf}} + 1)/2 < \bar{s}_{M2} = (\tilde{s}_r^{\text{lf}} + 1)/2$ yields

$$\mu(s) \equiv \frac{4\hat{c}}{s} - \tilde{s}_m^{\text{lf}} - \tilde{s}_r^{\text{lf}} \geq 0. \tag{A.125}$$

Since $\mu'(s) < 0$, it is sufficient to have $\mu(1) \geq 0$ for $\mu(s) \geq 0$. Using $\tilde{s}_m^{\text{lf}} = (1+4\hat{c} - \sqrt{1 + 16\hat{c}^2})/2$ from Eq. (5.26) and $\tilde{s}_r^{\text{lf}} = 4\hat{c}/(1 - 4\hat{c} + \sqrt{1 + 16\hat{c}^2})/2$ from Eq. (5.27) we find that $\mu(1) = 0$ and, hence, $\mu(s) \geq 0$. ■

A.29 Proof of Proposition 5.7.2

Using the definition of $\bar{s}_{\ell t}$ in Eq. (5.23) and substituting Y from Eq. (5.28) in Eq. (5.30) we can rewrite the social planner's objective function as:

$$W(\tilde{s}_m, \tilde{s}_r) = \begin{cases} A\left[1 + \tilde{s}_m(1 - \tilde{s}_m)\right]/2 - 2(1 - \tilde{s}_m)c & \text{for } (a), \\ A\tilde{s}_m\tilde{s}_r(\tilde{s}_r - \tilde{s}_m)/4 + \sum_k A\left[1 + \tilde{s}_k(1 - \tilde{s}_k)\right]/4 - (1 - \tilde{s}_k)c & \text{for } (b), \\ A\left[1 + \tilde{s}_m(1 - \tilde{s}_m)\right]/2 - (1 - \tilde{s}_m)c & \text{for } (c), \end{cases} \tag{A.126}$$

with $k = m, r$. The corresponding first order conditions then follow as

$$\frac{\partial W(\tilde{s}_m, \tilde{s}_r)}{\partial \tilde{s}_m} = \frac{A(1 - 2\tilde{s}_m)}{2} + 2c \overset{!}{=} 0 \quad \text{for } (a), \tag{A.127}$$

$$\frac{\partial W(\tilde{s}_m, \tilde{s}_r)}{\partial \tilde{s}_m} = \frac{A(1 - 2\tilde{s}_m + \tilde{s}_r^2 - 2\tilde{s}_m\tilde{s}_r)}{4} + c \overset{!}{=} 0 \quad \text{for } (b), \tag{A.128}$$

$$\frac{\partial W(\tilde{s}_m, \tilde{s}_r)}{\partial \tilde{s}_r} = \frac{A(1 - 2\tilde{s}_r - \tilde{s}_m^2 + 2\tilde{s}_m\tilde{s}_r)}{4} + c \overset{!}{=} 0 \quad \text{for } (b), \tag{A.129}$$

$$\frac{\partial W(\tilde{s}_m, \tilde{s}_r)}{\partial \tilde{s}_m} = \frac{A(1 - 2\tilde{s}_m)}{2} + c \overset{!}{=} 0 \quad \text{for } (c). \tag{A.130}$$

Since in the cases (a) and (c) the return margin is fixed to $\tilde{s}_r = \tilde{s}_m$ and $\tilde{s}_r = 1$, respectively, the social planner only has to choose the optimal emigration cutoff $\tilde{s}_m$, and it follows immediately that in both cases we have $\tilde{s}_m^{\text{sp}}(\hat{c}) = \frac{1}{2} + \tilde{s}_m^{\text{lf}}(\hat{c})$, where $\tilde{s}_m^{\text{lf}}(\hat{c})$ is defined as in Eq. (5.26). In the intermediate case (b) the emigration cutoff $\tilde{s}_m^{\mathrm{sp}}(\hat{c})$ and the return cutoff $\tilde{s}_r^{\mathrm{sp}}(\hat{c})$ follow as the joint solution to Eqs. (A.128) and (A.129). An explicit analytical solution to Eqs. (A.128) and (A.129) exists. However, since the resulting expressions for $\tilde{s}_m^{\mathrm{sp}}(\hat{c})$ and $\tilde{s}_r^{\mathrm{sp}}(\hat{c})$ are lengthy and tedious to derive, we do not report them here. We rather plot them directly as a function of the only exogenous variable $\hat{c}$ in Figure 5.13. Of course we thereby have to distinguish between the cases (a), (b), and (c), which as in the *laissez-faire* equilibrium are separated by the cost thresholds $\hat{c}_a^{\mathrm{sp}} \leq \hat{c}_b^{\mathrm{sp}} \leq \hat{c}_c^{\mathrm{sp}}$. Again the critical cost level $\hat{c}_c^{\mathrm{sp}}$ separates the high cost case (c) from an equilibrium without migration. Using $\tilde{s}_m^{\mathrm{sp}}(\hat{c}_c^{\mathrm{sp}}) = \frac{1}{2} + \tilde{s}_m^{\mathrm{lf}}(\hat{c}_c^{\mathrm{sp}}) \stackrel{!}{=} 1$ in combination with $\tilde{s}_m^{\mathrm{lf}}(\hat{c}) = \hat{c}$ from Eq. (5.26) for case (c) we find $\hat{c}_c^{\mathrm{sp}} = 1/2$. Similarly, when focusing on case (b) we find that $\tilde{s}_r^{\mathrm{sp}}(\hat{c}_b^{\mathrm{sp}}) \stackrel{!}{=} 1$ implies $\hat{c}_b^{\mathrm{sp}} \approx 1/20$. Consistently, we find that $W_b^{\mathrm{sp}}(\hat{c}_b^{\mathrm{sp}}) = W_c^{\mathrm{sp}}(\hat{c}_b^{\mathrm{sp}})$, while $W_b^{\mathrm{sp}}(\hat{c}) > W_c^{\mathrm{sp}}(\hat{c})$ for $\hat{c} < \hat{c}_b^{\mathrm{sp}}$, where $W_b^{\mathrm{sp}}(\hat{c})$ and $W_c^{\mathrm{sp}}(\hat{c})$ are the aggregate welfare levels resulting from the substitution of $\tilde{s}_m^{\mathrm{sp}}(\hat{c})$ and $\tilde{s}_r^{\mathrm{sp}}(\hat{c})$ for the cases (b) and (c) into $W(\tilde{s}_m, \tilde{s}_r)$ from Eq. (A.126). Finally, note that scenario (a) in principle is defined over $\hat{c} < 1/4$, which follows from $\tilde{s}_m^{\mathrm{sp}}(\hat{c}) = \frac{1}{2} + \tilde{s}_m^{\mathrm{lf}}(\hat{c}) < 1$ in combination with $\tilde{s}_m^{\mathrm{lf}}(\hat{c}) = 2\hat{c}$ from Eq. (5.26) for case (a). However, since $W_c^{\mathrm{sp}}(\hat{c}) > W_a^{\mathrm{sp}}(\hat{c})$ for all $\hat{c} < 1/4$ we have $\hat{c}_a^{\mathrm{sp}} = 0$. Taking stock, the aggregate welfare level in the social-planner equilibrium as depicted in Figure 5.13 then is defined as the envelope function $W^{\mathrm{sp}}(\hat{c}) = \max\{W_a^{\mathrm{sp}}(\hat{c}), W_b^{\mathrm{sp}}(\hat{c}), W_c^{\mathrm{sp}}(\hat{c})\}$. ■

A.30 Proof of Proposition 5.8.1

Suppose individuals now live for $t = 1, 2, 3$ periods. We focus on skilled migration and follow the baseline model in ruling out late emigration at age $t = 3$. Individual education and migration decisions can then be summarised in Figure A.2. The four possible individual outcomes in Figure A.2 imply an overall number of seven potential selection patterns for educated emigrants, which we denote by:

(d) $0 < \tilde{s}_e < \tilde{s}_m < \tilde{s}_r < 1 \;\Rightarrow$ *imperfect selection into temporary (permanent) migration*,

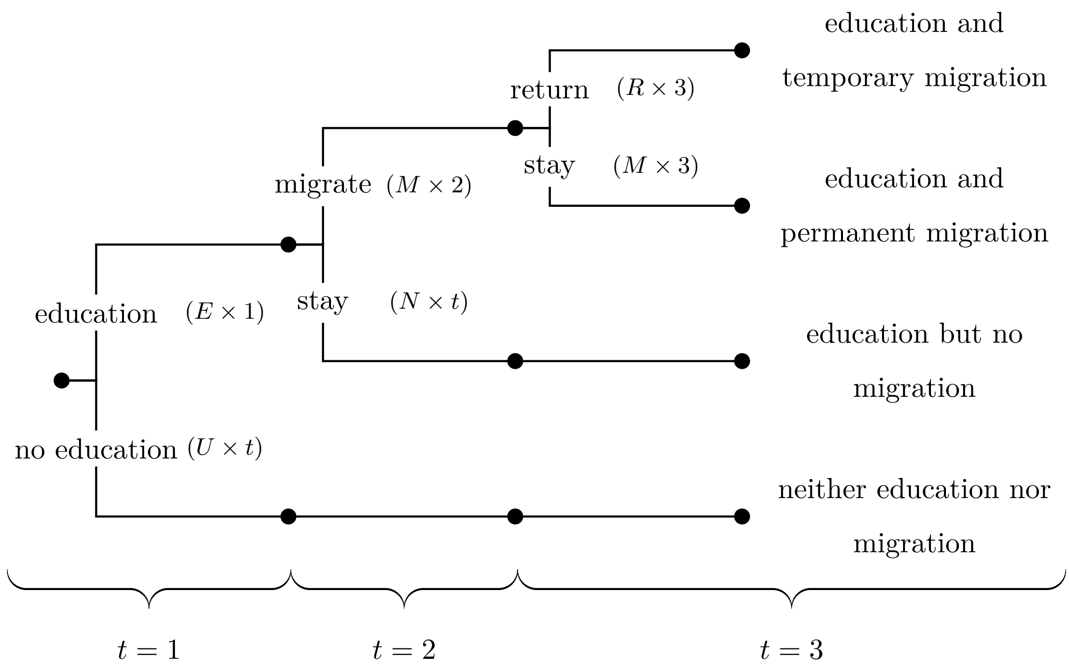
(e) $0 < \tilde{s}_e < \tilde{s}_m < \tilde{s}_r = 1 \;\Rightarrow$ *imperfect selection into temporary migration only*,

(f) $0 < \tilde{s}_e < \tilde{s}_m = \tilde{s}_r = 1 \;\Rightarrow$ *no selection into migration*,

(g) $0 < \tilde{s}_e = \tilde{s}_m = \tilde{s}_r < 1 \;\Rightarrow$ *perfect selection into permanent migration only*,

(h) $0 < \tilde{s}_e = \tilde{s}_m < \tilde{s}_r = 1 \;\Rightarrow$ *perfect selection into temporary migration only*,

Figure A.2: *Education, migration and return decisions*



(i) $0 < \tilde{s}_e = \tilde{s}_m < \tilde{s}_r < 1 \;\; \Rightarrow$ *(im)perfect selection into temporary (permanent) migration,*

(j) $0 < \tilde{s}_e < \tilde{s}_m = \tilde{s}_r < 1 \;\; \Rightarrow$ *(im)perfect selection into permanent migration only.*

From Section 5.6 we know that for constant and equal signalling cost $c_e = c_m = c > 0$ the signalling patterns (g) to (j) cannot exist. The only relevant scenarios are thus the cases (d), (e), and (f), which we discuss separately in the following. In scenario (f) only education is used as a signal, implying $0 < \tilde{s}_e < \tilde{s}_m = \tilde{s}_r = 1$. The critical student at age $t = 1$ hence has the choice between a career as an uneducated worker $(U \times t)$ or an educated $(E \times 1)$, who stays at home $(N \times t)$ at age $t = 2, 3$. The corresponding indifference condition for the marginal student $\tilde{s}_e$ thus reads:

$$\Delta_e\left(\tilde{s}_e\right) \equiv w\left(\bar{s}_{E1}, \tilde{s}_e\right) + w\left(\bar{s}_{N2}, \tilde{s}_e\right) + w\left(\bar{s}_{N3}, \tilde{s}_e\right) - c - \sum_{t=1}^{3} w\left(\bar{s}_{Ut}, \tilde{s}_e\right) \stackrel{!}{=} 0. \tag{A.131}$$

Using $w(\bar{s}_{\ell t}, s)$ from Eq. (5.20) and replacing $\bar{s}_{Ut} = \tilde{s}_e/2 < \bar{s}_{E1} = \bar{s}_{N2} = \bar{s}_{N3} = (\tilde{s}_e + 1)/2$ we find:

$$\tilde{s}_e^{\mathrm{lf}}\left(\hat{c}\right) = 2\hat{c}/3 < \tilde{s}_m^{\mathrm{lf}}\left(\hat{c}\right) = \tilde{s}_r^{\mathrm{lf}}\left(\hat{c}\right) = 1 \quad \text{for} \quad \hat{c} \in [\hat{c}_e^{\mathrm{lf}}, \hat{c}_f^{\mathrm{lf}}). \tag{A.132}$$

In scenario (e) the most high-skilled workers with $s \geq \tilde{s}_m$ combine education with subsequent (temporary) migration in order to obtain a more effective signalling device. We solve the resulting two-stage game by backward induction, beginning with the emigration decision at age $t = 2$. For the marginal emigrant $\tilde{s}_m$ at age $t = 2$ indifference condition:

$$\Delta_m(\tilde{s}_m) \equiv w(\bar{s}_{M2}, \tilde{s}_m) + w(\bar{s}_{R3}, \tilde{s}_m) - c - w(\bar{s}_{N2}, \tilde{s}_m) - w(\bar{s}_{N3}, \tilde{s}_m) \stackrel{!}{=} 0 \quad \text{(A.133)}$$

has to hold. Using the definition of $w(\bar{s}_{\ell t}, s)$ in Eq. (5.20) and replacing $\bar{s}_{Nt} = (\tilde{s}_e + \tilde{s}_m)/2 < \bar{s}_{M2} = \bar{s}_{R3} = (\tilde{s}_m + 1)/2$ we can solve for the emigration cutoff $\tilde{s}_m(\tilde{s}_e)$ as a function of the education cutoff $\tilde{s}_e$

$$\tilde{s}_m(\tilde{s}_e) = \frac{\hat{c}}{1 - \tilde{s}_e}. \quad \text{(A.134)}$$

The indifference condition for the marginal student at age $t = 3$ can be summarised as:

$$\Delta_e(\tilde{s}_e) \equiv w(\bar{s}_{E1}, \tilde{s}_e) + w(\bar{s}_{N2}, \tilde{s}_e) + w(\bar{s}_{N3}, \tilde{s}_e) - c - \sum_{t=1}^{3} w(\bar{s}_{Ut}, \tilde{s}_e) \stackrel{!}{=} 0. \quad \text{(A.135)}$$

Using the definition of $w(\bar{s}_{\ell t}, s)$ in Eq. (5.20) and replacing $\bar{s}_{Ut} = \tilde{s}_e/2 < \bar{s}_{Nt} = (\tilde{s}_e + \tilde{s}_m)/2 < \bar{s}_{E1} = (\tilde{s}_e + 1)/2$ we can rewrite this indifference condition as $\Delta_e(\tilde{s}_e) \equiv A(1 + 2\tilde{s}_m)\tilde{s}_e/2 - c \stackrel{!}{=} 0$, which solves for the education cutoff:

$$\tilde{s}_e^{\text{lf}}(\hat{c}) = \left(1 + 4\hat{c} - \sqrt{1 + 16\hat{c}^2}\right)/2 \quad \text{for} \quad \hat{c} \in [\hat{c}_d^{\text{lf}}, \hat{c}_e^{\text{lf}}), \quad \text{(A.136)}$$

once $\tilde{s}_m$ is replaced by $\tilde{s}_m(\tilde{s}_e) = \hat{c}/(1 - \tilde{s}_e)$ from Eq. (A.134). Substituting $\tilde{s}_e^{\text{lf}}(\hat{c})$ from Eq. (A.136) back into $\tilde{s}_m(\tilde{s}_e) = \hat{c}/(1 - \tilde{s}_e)$ from Eq. (A.134) then yields the corresponding migration cutoff:

$$\tilde{s}_m^{\text{lf}}(\hat{c}) = 2\hat{c}/\left(1 + 4\hat{c} - \sqrt{1 + 16\hat{c}^2}\right) \quad \text{for} \quad \hat{c} \in [\hat{c}_d^{\text{lf}}, \hat{c}_e^{\text{lf}}). \quad \text{(A.137)}$$

Finally, in case (d) with $0 < \tilde{s}_e < \tilde{s}_m < \tilde{s}_r < 1$ all workers with skills $s \geq \tilde{s}_e$ get educated before the most high skilled among them with $s \geq \tilde{s}_m$ ($s \geq \tilde{s}_r$) emigrate temporary (permanently) abroad. We begin with the return decision of initial emigrants at age $t = 3$ and solve by backward induction. At age $t = 3$ worker $\tilde{s}_r$ is indifferent between returning home ($R \times 3$) and staying abroad ($M \times 3$) if

$$\Delta_r(\tilde{s}_r) \equiv w(\bar{s}_{M3}, \tilde{s}_r) - c - w(\bar{s}_{R3}, \tilde{s}_r) \stackrel{!}{=} 0, \quad \text{(A.138)}$$

whereas $\bar{s}_{R3} = (\tilde{s}_m + \tilde{s}_r)/2 < \bar{s}_{M3} = (\tilde{s}_r + 1)/2$. Using the definition of $w(\bar{s}_{\ell t}, s)$ in Eq. (5.20) we can solve for

$$\tilde{s}_r(\tilde{s}_m) = \frac{2\hat{c}}{1 - \tilde{s}_m}. \tag{A.139}$$

At age $t = 2$ worker $\tilde{s}_m$ is indifferent between going abroad ($M \times 2$) and staying at home ($N \times 2$) if

$$\Delta_m(\tilde{s}_m) \equiv w(\bar{s}_{M2}, \tilde{s}_m) + w(\bar{s}_{R3}, \tilde{s}_m) - c - w(\bar{s}_{N2}, \tilde{s}_m) - w(\bar{s}_{N3}, \tilde{s}_m) \overset{!}{=} 0, \tag{A.140}$$

whereas $\bar{s}_{Nt} = (\tilde{s}_e + \tilde{s}_m)/2 < \bar{s}_{R3} = (\tilde{s}_m + \tilde{s}_r)/2 < \bar{s}_{M2} = (\tilde{s}_m + 1)/2$. Using the definition of $w(\bar{s}_{\ell t}, s)$ in Eq. (5.20) and replacing $\tilde{s}_r$ by $\tilde{s}_r(\tilde{s}_m)$ from Eq. (A.139) we can solve for

$$\tilde{s}_m = \frac{1 - 2\tilde{s}_e + 4\hat{c} - \sqrt{1 + 16\hat{c}^2 - 4\tilde{s}_e(1 - \tilde{s}_e)}}{2(1 - 2\tilde{s}_e)}. \tag{A.141}$$

Finally, at age $t = 1$ worker $\tilde{s}_e$ is indifferent between getting educated ($E \times 1$) and staying uneducated ($U \times 1$) if

$$\Delta_e(\tilde{s}_e) \equiv w(\bar{s}_{E1}, \tilde{s}_e) + w(\bar{s}_{N2}, \tilde{s}_e) + w(\bar{s}_{N3}, \tilde{s}_e) - c - \sum_{t=1}^{3} w(\bar{s}_{Ut}, \tilde{s}_e) \overset{!}{=} 0, \tag{A.142}$$

whereas $\bar{s}_{Ut} = \tilde{s}_e/2 < \bar{s}_{Nt} = (\tilde{s}_e + \tilde{s}_m)/2 < \bar{s}_{E1} = (\tilde{s}_e + 1)/2$. Using the definition of $w(\bar{s}_{\ell t}, s)$ in Eq. (5.20) and replacing $\tilde{s}_m$ by $\tilde{s}_m(\tilde{s}_e)$ Eq. (A.141) we can solve for $\tilde{s}_e^{\text{lf}}(\hat{c})$. Substituting $\tilde{s}_e^{\text{lf}}(\hat{c})$ back into the Eq. (A.141) then delivers $\tilde{s}_m^{\text{lf}}(\hat{c})$. Once obtained, $\tilde{s}_m^{\text{lf}}(\hat{c})$ from Eq. (A.141) can then be used to replace $\tilde{s}_m$ in $\tilde{s}_r(\tilde{s}_m)$ from Eq. (A.139), which finally results in $\tilde{s}_r^{\text{lf}}(\hat{c})$. Analytical solutions for $\tilde{s}_e^{\text{lf}}(\hat{c})$, $\tilde{s}_m^{\text{lf}}(\hat{c})$, and $\tilde{s}_r^{\text{lf}}(\hat{c})$ exist, but are lengthy and tedious to derive. We, thus, instead of reporting the cutoffs $\tilde{s}_e^{\text{lf}}(\hat{c})$, $\tilde{s}_m^{\text{lf}}(\hat{c})$, and $\tilde{s}_r^{\text{lf}}(\hat{c})$ here directly, plot them as functions of the only exogenous variable $\hat{c} > 0$ in Figure 5.15. The cost thresholds $0 < \hat{c}_d^{\text{lf}} < \hat{c}_e^{\text{lf}} < \hat{c}_f^{\text{lf}}$, separating the scenarios (d), (e), and (f), thereby follow directly from the definition of the respective cases. Using $\tilde{s}_e^{\text{lf}}(\hat{c})$ from Eq. (A.132) we find that $\tilde{s}_e^{\text{lf}}(\hat{c}_f^{\text{lf}}) = 2\hat{c}_f^{\text{lf}}/3 \overset{!}{=} 1$ implies $\hat{c}_f^{\text{lf}} = 3/2$. Similarly, $\tilde{s}_m^{\text{lf}}(\hat{c}_e^{\text{lf}}) \overset{!}{=} 1$ with $\tilde{s}_m^{\text{lf}}(\hat{c})$ from Eq. (A.137) yields $\hat{c}_e^{\text{lf}} = 3/5$. Finally, $\tilde{s}_r^{\text{lf}}(\hat{c}_d^{\text{lf}}) \overset{!}{=} 1$ with $\tilde{s}_r^{\text{lf}}(\hat{c})$ for scenario (d) solves for $\hat{c}_d^{\text{lf}} \approx 3/10$. ■

A.31 Derivation of Eqs. (5.26″) and (5.27″)

Note that discounting has no effect on individuals' return decision at age $t = 2$. Hence, as in the baseline model, the return cutoff $\tilde{s}_r(\tilde{s}_m)$ is linked to the emigration cutoff $\tilde{s}_r$ through Eq. (5.22).

The emigration decision at age $t = 1$, on the contrary, crucially depends on how individuals value future costs and benefits. Modifying the indifference condition in Eq. (5.24) accordingly we find that the net gain from emigration can be written as:

$$\Delta_m(\tilde{s}_m) = \begin{cases} w(\bar{s}_{M1}, \tilde{s}_m) + \delta w(\bar{s}_{M2}, \tilde{s}_m) - (1+\delta)c - w(\bar{s}_{N1}, \tilde{s}_m) - \delta w(\bar{s}_{N2}, \tilde{s}_m) & (a), \\ w(\bar{s}_{M1}, \tilde{s}_m) + \delta w(\bar{s}_{R2}, \tilde{s}_m) - c - w(\bar{s}_{N1}, \tilde{s}_m) - \delta w(\bar{s}_{N2}, \tilde{s}_m) & (b), \\ w(\bar{s}_{M1}, \tilde{s}_m) + \delta w(\bar{s}_{R2}, \tilde{s}_m) - c - w(\bar{s}_{N1}, \tilde{s}_m) - \delta w(\bar{s}_{N2}, \tilde{s}_m) & (c), \end{cases} \tag{A.143}$$

whereas $\Delta_m(\tilde{s}_m) \stackrel{!}{=} 0$ has to hold for the critical emigrant $\tilde{s}_m$. Using the definition of $w(\bar{s}_{\ell t}, s)$ in Eq. (5.20) in combination with $\bar{s}_{\ell t}\ \forall\ \ell \in \{N, M, R\}$ and $t = 1, 2$ from Eq. (5.23) then allows us to recover $\tilde{s}_m^{\text{lf}}(\hat{c}, \delta)$ in Eq. (5.26″) separately for the cases (*a*), (*b*), and (*c*). Replacing $\tilde{s}_m$ in $\tilde{s}_r(\tilde{s}_m) = 2\hat{c}/(1 - \tilde{s}_m)$ from Eq. (5.22) by $\tilde{s}_m^{\text{lf}}(\hat{c}, \delta)$ from Eq. (5.26″) then finally yields $\tilde{s}_m^{\text{lf}}(\hat{c}, \delta)$ as given in Eq. (5.27″). ■

A.32 Proof of Proposition 6.4.1

The indirect utility of Home's median voter $V(p, \tilde{s}, \breve{s}^*)$ is given by

$$\begin{aligned} V(p, \tilde{s}, \breve{s}^*) = &\frac{1}{P(p,1)} \\ &- \frac{p}{P(p,1)} A \left[\frac{\int_{\tilde{s}}^1 f(s)\, ds L}{\int_{\tilde{s}}^1 f(s)\, ds L + \int_{\breve{s}^*}^1 f(s^*)\, ds^* L^*} \beta \int_{\tilde{s}}^1 (s - \tilde{s}) f(s)\, ds \right] \\ &- \frac{p}{P(p,1)} A \left[\frac{\int_{\breve{s}^*}^1 f(s^*)\, ds^* L^*}{\int_{\tilde{s}}^1 f(s)\, ds L + \int_{\breve{s}^*}^1 f(s^*)\, ds^* L^*} \beta \int_{\breve{s}^*}^1 (s^* - \tilde{s}) f(s^*)\, ds^* \right], \end{aligned} \tag{A.144}$$

in which $P(p, 1)$ denotes the aggregate price index, which is increasing in p. The first term at the right hand side of (A.144) reflects real income, the second term is disutility from comparing to other natives, $\int_{\tilde{s}}^1 f(s)\, ds L$, while the last term measures disutility from comparing to migrants, $\int_{\breve{s}^*}^1 f(s^*)\, ds^* L^*$. The median voter supports migration if $V(p^m, \tilde{s}^m, \breve{s}^*) > V(p^a, \tilde{s}^a, 1)$. Assuming $f(s) = f(s^*)1$ this inequality can be expressed as:

$$\frac{1}{P(p^a,1)} - \frac{p^a}{P(p^a,1)} \beta \int_{\tilde{s}^a}^1 (s - \tilde{s}^a) f(s)\, ds < \frac{1}{P(p^m,1)} - \frac{p^m}{P(p^m,1)} \beta \tag{A.145}$$

$$\left[\frac{(1 - \tilde{s}^m) L}{(1 - \tilde{s}^m) L + (1 - \breve{s}^m) L^*} \int_{\tilde{s}^m}^1 s - \tilde{s}^m ds + \frac{(1 - \breve{s}^m) L^*}{(1 - \tilde{s}^m) L + (1 - \breve{s}^m) L^*} \int_{\breve{s}^m}^1 s - \tilde{s}^m ds \right], \tag{A.146}$$

which simplifies to $\frac{1}{2}(1-\tilde{s}^a)^2 > \frac{1}{2}[1-(\breve{s}^*)^2]-(1-\breve{s}^*)\,\tilde{s}^m$. Defining the right hand side of this inequality as $\xi(\breve{s}^*) \equiv \frac{1}{2}[1-(\breve{s}^*)^2]-(1-\breve{s}^*)\,\tilde{s}^m$ and noting that $\xi'(\breve{s}^*) = \tilde{s}^m - \breve{s}^* < 0$ we get $\frac{1}{2}(1-\tilde{s}^a)^2 > \frac{1}{2}(1-\tilde{s}^m)^2 > \frac{1}{2}[1-(\breve{s}^*)]-(1-\breve{s}^*)\,\tilde{s}^m$ which holds for all $\tilde{s}^m < \breve{s}^*$ as $\tilde{s}^a < \tilde{s}^m$. ∎

B Empirical appendix

B.1 Data description and summary statistics

We source our data from the following providers: The data on (nominal) output at the industry level stem from the OECD's STAN database. Deflation of output for the calculation of the real growth rate control is done with industry specific producer price indices obtained from the German Statistical Office. Industry competition for 2005 is taken from the Monopoly Commission's annual report to the Federal German government and can be accessed at `http://www.monopolkommission.de/haupt.html`. The input output tables used in the calculation of the offshoring indices are part of the national accounts provided by the German Statistical Office at `https://www.destatis.de/EN/Homepage.html`. Data on industry level trade and output are taken from the OECD STAN data base, as are the export shares in production.

Table 8.1 summarizes all variables in our final sample of 3,917 individuals which are full-time employed in manufacturing. More than half (59%) of the respondents participated in on-the-job training. Of those who participated, 42% did so by own initiative. We can group individuals into five age groups, with the average worker being of age 42 having 14 years of tenure. Unsurprisingly, the majority of workers (76%) in manufacturing are male. We classify workers according to their education as high-skilled (university degree), medium-skilled (degree from a technical school, e.g. the German "Meister") and low-skilled (all residual workers). The majority of workers (68%) are classified as low-skilled, less are high- (21%) or medium-skilled (11%). Of the respondents 17% stated that having a career is important for them. Only a small fraction of all workers held a fixed term contract (6%) or were just temporarily employed (1%). Among all workers 10 % answered that they face the fear of job loss. We classify employers according to the number of employees and distinguish between five groups: firms with 1 - 9, 10 - 49, 50 - 249, 250 - 499 and more than 500 employees. The majority of firms (90%) introduced new technologies

Table 8.1: *Summary statistics: estimation sample*

Variables	share	mean	st. dev.
Individual characteristics			
On-the-job training	0.586	-	-
Thereof by own initiative	0.421	-	-
Age	-	42.06	10.06
16 - 29	0.117	-	-
30 - 39	0.297	-	-
40 - 49	0.345	-	-
50 - 64	0.231	-	-
≥ 65	0.010	-	-
Education			
Low	0.677	-	-
Medium	0.109	-	-
High	0.214	-	-
Further individual characteristics			
Important to have a career	0.173	-	-
Fixed term contract	0.055	-	-
Temporary work	0.011	-	-
Job loss fear	0.103	-	-
Female	0.241	-	-
Tenure		13.66	9.621
Employer size (# of employees)			
1 - 9	0.109	-	-
10 - 49	0.183	-	-
50 - 249	0.246	-	-
250 - 499	0.132	-	-
≥ 500	0.332	-	-
Further employer characteristics			
New technology introduced	0.896	-	-
Current firm success (very) good	0.805	-	-
Industry characteristics			
Offshoring growth 2004 - 2006	-	.325	0.393
Output growth 2004 - 2006	-	0.079	0.114
Herfindahl index (x 1,000) 2005	-	60.363	83.684
Number of observations	3,917		

during the sample period. Overall the employing firm's success was largely seen as good or very good; 81% of the respondents answered in this way. Industry output growth is the growth of real output, calculated as log-difference. The Herfindahl index of industry concentration is the sum of the squared market shares of all market participants in the respective 2-digit NACE 1.1 industry.

Table 8.2: *Summary statistics: offshoring*

j	Industry classification	O_j	$\widehat{O}_j$	j	Industry classification	O_j	$\widehat{O}_j$
35	Other transport equip.	0.82	152.65	22	Publishing, printing	0.047	26.68
34	Motor vehicles	0.37	95.47	30	Office, computing mach.	8.17	23.76
27	Basic metals	2.15	86.07	15	Food, beverages	0.54	22.88
33	Medical, optical, precision instr.	0.61	52.01	29	Machinery, equipment	1.71	20.63
28	Fabricated metal prod.	0.30	39.38	20	Wood, cork prod.	1.02	19.23
25	Rubber, plastic	0.16	34.81	26	Non-metallic mineral prod.	0.31	13.31
24	Chemicals	0.83	34.32	36	Furniture	3.20	10.49
16	Tobacco	0.11	31.67	17	Textiles	4.63	8.03
18	Wearing apparel	5.60	30.45	32	Radio, television, comm.	9.83	5.36
19	Leather, luggage	7.70	29.27	31	Electrical machinery	1.52	4.19
21	Paper	0.49	27.79	23	Coke, refined petroleum	0.456	-52.44

Notes: The offshoring intensity O_j (in percent) is calculated for 2004. Offshoring growth $\widehat{O}_j$ (in percent) is calculated over the time span from 2004 to 2006. Industries are ranked in decreasing order according to the magnitude of sectoral offshoring growth.

Industry level offshoring is calculated as described in Eq. (4.7). For the industries 15-16, 17-19, and 21-22 the OECD STAN bilateral trade data base only holds information on combined non-OECD trade flows. We hence use the same share of non-OECD imports in total imports for the individual industries within each of the three aggregates and multiply them with total STAN imports, for which we have industry specific data in all cases. Checking the robustness of this approach, we dropped the respective sectors and still found our results presented in section 4.2.3 to be similarly sized and statistically significant. Table 8.2 gives an overview of offshoring intensities across industries, both in levels and growth rates.

Table 8.3: *Offshoring and on-the-job training: robustness*

	(1) import pen.	(2) with exports	(3) weighted	(4) no small ind.	(5) no 35 and 32	(6) no 24 no 17	(7) own initiative
Average marginal effect of:							
Offshoring growth	0.2423***	0.0815***	0.0626***	0.0748***	0.0850***	0.0827***	0.0763***
	(0.0664)	(0.0214)	(0.0192)	(0.0245)	(0.0229)	(0.0191)	(0.0203)
Age 30 - 39	-0.0045	-0.0043	-0.0306	-0.0016	-0.0033	0.0058	0.0171
	(0.0196)	(0.0196)	(0.0235)	(0.0203)	(0.0205)	(0.0196)	(0.0287)
Age 40 - 49	-0.0505**	-0.0501**	-0.0599**	-0.0462*	-0.0475**	-0.0420*	-0.0091
	(0.0231)	(0.0233)	(0.0296)	(0.0241)	(0.0234)	(0.0236)	(0.0308)
Age 50 - 64	-0.1406***	-0.1380***	-0.1602***	-0.1376***	-0.1318***	-0.1346***	-0.0875**
	(0.0278)	(0.0282)	(0.0327)	(0.0288)	(0.0289)	(0.0296)	(0.0362)
Age 65+	-0.3362***	-0.3338***	-0.3759***	-0.3343***	-0.3394***	-0.3256***	-0.1713***
	(0.0573)	(0.0568)	(0.0618)	(0.0577)	(0.0620)	(0.0616)	(0.0558)
Female	-0.0678***	-0.0678***	-0.1032***	-0.0645***	-0.0687***	-0.0655***	-0.0468**
	(0.0190)	(0.0189)	(0.0198)	(0.0187)	(0.0205)	(0.0222)	(0.0194)
Married	-0.0103	-0.0104	-0.0068	-0.0129	-0.0114	-0.0012	-0.0092
	(0.0215)	(0.0214)	(0.0243)	(0.0214)	(0.0226)	(0.0242)	(0.0227)
Tenure	0.0028	0.0028	0.0040	0.0027	0.0029	0.0008	-0.0015
	(0.0040)	(0.0039)	(0.0042)	(0.0040)	(0.0042)	(0.0032)	(0.0051)
Tenure squared	-0.0000	-0.0000	-0.0000	-0.0000	-0.0001	0.0000	0.0000
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Medium-skill	0.0363	0.0346	0.0246	0.0343	0.0308	0.0221	0.1134***
	(0.0353)	(0.0344)	(0.0336)	(0.0351)	(0.0359)	(0.0353)	(0.0411)
High-skill	-0.0002	0.0026	-0.0295	-0.0026	0.0006	0.0178	0.0253
	(0.0213)	(0.0216)	(0.0249)	(0.0218)	(0.0234)	(0.0213)	(0.0242)
Importance to have a career	0.0655***	0.0633***	0.0632***	0.0613***	0.0601***	0.0568***	0.0808***
	(0.0208)	(0.0204)	(0.0235)	(0.0207)	(0.0209)	(0.0215)	(0.0203)
Firm size 10 - 49	-0.0167	-0.0181	-0.0029	-0.0239	-0.0148	-0.0192	-0.1247***
	(0.0212)	(0.0212)	(0.0245)	(0.0218)	(0.0212)	(0.0219)	(0.0360)
Firm size 50 - 249	0.0533***	0.0505***	0.0690***	0.0488***	0.0546***	0.0401***	-0.0561***
	(0.0168)	(0.0162)	(0.0259)	(0.0162)	(0.0155)	(0.0151)	(0.0217)
Firm size 250 - 499	0.1021***	0.0977***	0.0874**	0.0914***	0.1014***	0.1007***	-0.0260
	(0.0290)	(0.0277)	(0.0387)	(0.0289)	(0.0288)	(0.0304)	(0.0243)
Firm size 500+	0.1290***	0.1203***	0.1041**	0.1152***	0.1115***	0.1119***	0.0309
	(0.0234)	(0.0241)	(0.0464)	(0.0239)	(0.0233)	(0.0272)	(0.0351)
Fixed term contract	-0.0790**	-0.0794**	-0.0637	-0.0801**	-0.0788**	-0.0993***	-0.0985***
	(0.0342)	(0.0330)	(0.0456)	(0.0331)	(0.0343)	(0.0334)	(0.0331)
Temporary work	0.0392	0.0418	0.0506	0.0156	0.0680	0.0445	0.0375
	(0.0522)	(0.0527)	(0.0932)	(0.0574)	(0.0534)	(0.0545)	(0.0957)
Job loss fear	-0.0517**	-0.0511**	-0.0616**	-0.0525**	-0.0562***	-0.0481**	-0.0426
	(0.0215)	(0.0209)	(0.0299)	(0.0212)	(0.0208)	(0.0210)	(0.0327)
New technology introduced	0.1637***	0.1646***	0.1488***	0.1633***	0.1687***	0.1676***	0.1293***
	(0.0220)	(0.0215)	(0.0269)	(0.0218)	(0.0224)	(0.0221)	(0.0285)
Current firm success (very) good	0.0341*	0.0393**	0.0401	0.0396**	0.0446**	0.0405*	0.0471**
	(0.0203)	(0.0198)	(0.0246)	(0.0193)	(0.0204)	(0.0210)	(0.0188)
Herfindahl index	0.0010***	0.0007***	0.0008***	0.0007***	0.0006***	0.0005***	0.0004**
	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0001)	(0.0001)	(0.0002)
Growth in export share of prod.		0.1904					
		(0.1495)					
KldB88 (2-digit) occupation FE	yes	yes	yes	yes	yes	yes	yes
Pseudo R-squared	0.1397	0.1395	0.1383	0.1393	0.1393	0.1279	0.2207
Observations	3888	3888	3888	3,845	3,685	3,421	2,623

Notes: The table shows average marginal effects from estimating variants of the Probit model specified in section 4.2.1. The reference category for firm size is 1 - 9 employees. The industry output growth is compute for 2004 to 2006. The Herfindahl index, which is published bi-annually by the German Monopoly Commission refers to 2005. Individual controls are the same as in column (6) of Table 4.1. Industry level controls are as in Table 4.2. Export share in production is taken from the OECD STAN data base. Standard errors are clustered at the industry level and shown in parentheses below the coefficients. Superscripts ***, **, and * denote statistical significance at the 1%, 5%, and 10% level, respectively.

Bibliography

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, vol. 4, pp. 1043–1171. Elsevier, North Holland/NL.

Acemoglu, D., G. Gancia, and F. Zilibotti (2012): "Offshoring and Directed Technical Change," CEPR Discussion Paper 9247.

Acemoglu, D., and J.-S. Pischke (1998): "Why Do Firms Train? Theory and Evidence," Quarterly Journal of Economics, 113(1), 78–118.

Akerlof, G. A. (1982): "Labor Contracts as Partial Gift Exchange," Quarterly Journal of Economics, 97(4), 543–569.

Akerlof, G. A., and J. L. Yellen (1990): "The Fair Wage-Effort Hypothesis and Unemployment," Quarterly Journal of Economics, 105(2), 255–283.

Alpizar, F., F. Carlsson, and O. Johansson-Stenman (2005): "How Much Do We Care about Absolute Versus Relative Income and Consumption?," Journal of Economic Behavior & Organization, 56(3), 405 – 421.

Amiti, M., and D. R. Davis (2012): "Trade, Firms, and Wages: Theory and Evidence," Review of Economic Studies, 79(1), 1–36.

Andersson, F., M. García-Pérez, J. C. Haltiwanger, K. McCue, and S. Sanders (2010): "Workplace Concentration of Immigrants," NBER Working Papers 16544.

Antràs, P., L. Garicano, and E. Rossi-Hansberg (2006): "Offshoring in a Knowledge Economy," Quarterly Journal of Economics, 121(1), 31–77.

Antràs, P., and E. Helpman (2004): "Global Sourcing," Journal of Political Economy, 112(3), 552–580.

Antràs, P., and E. Helpman (2008): "Contractual Frictions and Global Sourcing," in The Organization of Firms in a Global Economy, ed. by E. Helpman, D. Marin, and T. Verdier. Harvard University Press, Cambridge/MA.

Ardelean, A. (2011): "How Strong Is Love of Variety," Unpublished Manuscript, University of Santa Clara.

Arkolakis, C. (2010): "Market Penetration Costs and the New Consumers Margin in International Trade," Journal of Political Economy, 118(6), 1151–1199.

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

Artuç, E., and J. McLaren (2012): "Trade Policy and Wage Inequality: A Structural Analysis with Occupational and Sectoral Mobility," Unpublished Manuscript, University of Virginia.

Arulampalam, W., A. L. Booth, and M. L. Bryan (2004): "Training in Europe," Journal of the European Economic Association, 2(2-3), 346–360.

Aslund, O., and O. N. Skans (2010): "Will I See You at Work? Ethnic Workplace Segregation in Sweden, 1985-2002," Industrial and Labor Relations Review, 63(3), 471–493.

Aydemir, A., and C. Robinson (2008): "Global Labour Markets, Return, and Onward Migration," Canadian Journal of Economics, 41(4), 1285–1311.

Baldwin, R., and J. Lopez-Gonzalez (2013): "Supply-Chain Trade: A Portrait of Global Patterns and Several Testable Hypotheses," NBER Working Papers 18957, National Bureau of Economic Research, Inc.

BALDWIN, R., AND A. J. VENABLES (2013): "Spiders and Snakes: Offshoring and Agglomeration in the Global Economy," Journal of International Economics, 90(2), 245–254.

BARRETT, A., AND J. GOGGIN (2010): "Returning to the Question of a Wage Premium for Returning Migrants," National Institute Economic Review, 213(1), R43–R51.

BARRETT, A., AND P. J. O'CONNELL (2001): "Is There a Wage Premium for Returning Irish Migrants?," The Economic and Social Review, 32(1), 1–21.

BASSANINI, A., A. BOOTH, G. BRUNELLO, M. DE PAOLA, AND E. LEUVEN (2007): Education and Training in Europe, chap. Workplace Training in Europe, pp. 143 – 309. Oxford University Press.

BASSANINI, A., AND G. BRUNELLO (2011): "Barriers to Entry, Deregulation and Workplace Training: A Theoretical Model with Evidence from Europe," European Economic Review, 55(8), 1152–1176.

BASU, K. (1997): Analytical Development Economics: The Less Developed Economy Revisited. MIT Press, Cambridge, Mass.

BAUMGARTEN, D., I. GEISHECKER, AND H. GÖRG (2013): "Offshoring, Tasks, and the Skill-wage Pattern," European Economic Review, 61(0), 132 – 152.

BECKER, G. S. (1991): A Treatise on the Family. Harvard Univ. Pr., Cambridge, Mass.

BECKER, S. O., K. EKHOLM, AND M.-A. MUENDLER (2013): "Offshoring and the Onshore Composition of Tasks and Skills," Journal of International Economics, 90(1), 91–106.

BEINE, M., F. DOCQUIER, AND H. RAPOPORT (2001): "Brain Drain and Economic Growth: Theory and Evidence," Journal of Development Economics, 64(1), 275–289.

BENHABIB, J., AND B. JOVANOVIC (2012): "Optimal Migration: A World Perspective," International Economic Review, 53(2), 321–348.

BERGLANN, H., E. R. MOEN, K. RØED, AND J. F. SKOGSTRØM (2011): "Entrepreneurship: Origins and Returns," Labour Economics, 18(2), 180–193.

Bernard, A. B., J. B. Jensen, S. J. Redding, and P. K. Schott (2007): "Firms in International Trade," Journal of Economic Perspectives, 21(3), 105–130.

——— (2012): "The Empirics of Firm Heterogeneity and International Trade," Annual Review of Economics, 4(1), 283–313.

Berninghaus, S., and H. Seifert-Vogt (1988): "Temporary vs. Permanent Migration: A Decision Theoretical Approach," Journal of Population Economics, 1(3), 195–211.

Berninghaus, S., and H. G. Seifert-Vogt (1993): "The Role of the Target Saving Motive in Guest Worker Migration: A Theoretical Study," Journal of Economic Dynamics and Control, 17(1-2), 181–205.

Berry, R. A., and R. Soligo (1969): "Some Welfare Aspects of International Migration," Journal of Political Economy, 77(5), 778–794.

Bhagwati, J., and K. Hamada (1974): "The Brain Drain, International Integration of Markets for Professionals and Unemployment: A Theoretical Analysis," Journal of Development Economics, 1(1), 19–42.

Bhagwati, J., and C. Rodriguez (1975): "Welfare-theoretical Analyses of the Brain Drain," Journal of Development Economics, 2(3), 195–221.

Biswas, T., and J. McHardy (2005): "Measuring the Balance of Intra-regional Migration," Applied Economics, 37(19), 2221–2230.

Blanchflower, D. G., A. J. Oswald, and P. Sanfey (1996): "Wages, Profits, and Rent-Sharing," Quarterly Journal of Economics, 111(1), 227–251.

Blinder, A. S. (2006): "Offshoring: The Next Industrial Revolution?," Foreign Affairs, 85(2), 113–128.

——— (2009): "How Many US Jobs Might be Offshorable?," World Economics, 10(2), 41–78.

Blinder, A. S., and A. B. Krueger (2013): "Alternative Measures of Offshorability: A Survey Approach," Journal of Labor Economics, 31(2), S97–S128.

Borjas, G. J. (1987): "Self-Selection and the Earnings of Immigrants," American Economic Review, 77, 531–53.

Borjas, G. J. (1991): "Immigration and Self-selection," in Immigration, Trade and the Labor Market, ed. by J. M. Abowd, and R. B. Freeman, pp. 29–76. University of Chicago Press.

——— (1999): "The Economic Analysis of Immigration," in Handbook of Labor Economics, ed. by O. C. Ashenfelter, and D. Card, vol. 3, Part A, pp. 1697–1760. Elsevier.

Borjas, G. J., and B. Bratsberg (1996): "Who Leaves? The Outmigration of the Foreign-Born," The Review of Economics and Statistics, 78(1), 165–176.

Bougheas, S., and D. Nelson (2013): "On the Political Economy of High Skilled Migration and International Trade," European Economic Review, 63(C), 206–224.

Bougheas, S., and D. R. Nelson (2012): "Skilled Worker Migration and Trade: Inequality and Welfare," The World Economy, 35(2), 197–215.

Bougheas, S., and R. Riezman (2007): "Trade and the Distribution of Human Capital," Journal of International Economics, 73(2), 421–433.

Brülhart, M. (2009): "An Account of Global Intra-industry Trade, 1962-2006," The World Economy, 32(3), 401–459.

Broda, C., and D. E. Weinstein (2006): "Globalization and the Gains from Variety," Quarterly Journal of Economics, 121(2), 541–585.

Budd, J. W., J. Konings, and M. J. Slaughter (2005): "Wages and International Rent Sharing in Multinational Firms," Review of Economics and Statistics, 87(1), 73–84.

Caliendo, M., and S. Künn (2011): "Start-up Subsidies for the Unemployed: Long-term Evidence and Effect Heterogeneity," Journal of Public Economics, 95(3-4), 311–331.

Charness, G., and B. Grosskopf (2001): "Relative Payoffs and Happiness: An Experimental Study," Journal of Economic Behavior & Organization, 45(3), 301–328.

Chiswick, B. R., and P. W. Miller (2009): "The International Transferability of Immigrants' Human Capital," Economics of Education Review, 28(2), 162–169.

Co, C. Y., I. N. Gang, and M.-S. Yun (2000): "Returns to Returning," Journal of Population Economics, 13(1), 57–79.

Cohn, A., E. Fehr, and L. Goette (2013): "Fair Wages and Effort Provision: Combining Evidence from the Lab and the Field," ECON - Working Papers 107, Department of Economics - University of Zurich.

Corcos, G., M. D. Gatto, G. Mion, and G. I. Ottaviano (2012): "Productivity and Firm Selection: Quantifying the 'New' Gains from Trade," Economic Journal, 122(561), 754–798.

Costinot, A., J. Vogel, and S. Wang (2012): "Global Supply Chains and Wage Inequality," American Economic Review, Papers and Proceedings, 102(3), 396–401.

——— (2013): "An Elementary Theory of Global Supply Chains," Review of Economic Studies, 80(1), 109–144.

Crinò, R. (2008): "Offshoring, Multinationals and Labour Market: A Review of the Empirical Literature," Journal of Economic Surveys, 23(2), 197–249.

Crinò, R. (2010): "The Effects of Offshoring on Post-displacement Wages: Evidence from the United States," The World Economy, 33(12), 1836–1869.

Davidson, C., and S. J. Matusz (2006): "Trade Liberalization and Compensation," International Economic Review, 47(3), 723–747.

Davidson, C., S. J. Matusz, and D. R. Nelson (2007): "Can Compensation Save Free Trade?," Journal of International Economics, 71(1), 167–186.

Davidson, C., S. J. Matusz, and A. Shevchenko (2008): "Outsourcing Peter to Pay Paul: High-Skill Expectations and Low-Skill Wages with Imperfect Labor Markets," Macroeconomic Dynamics, 12(04), 463–479.

Davies, R., and R. Desbordes (2012): "Greenfield FDI and Skill Upgrading," CEPR Discussion Paper, (8912).

Defoort, C., and G. Rogers (2008): "Long-term Trends in International Migration: An Analysis of the Six Main Receiving Countries," Population (English Edition, 2002-), 63(2), pp. 285–317.

Dhingra, S., and J. Morrow (2013): "Monopolistic Competition and Optimum Product Diversity under Firm Heterogeneity," Unpublished Manuscript, London School of Economics.

Djajic, S. (1989): "Migrants in a Guest-worker System: A Utility Maximizing Approach," Journal of Development Economics, 31(2), 327–339.

——— (2010): "Investment Opportunities in the Source Country and Temporary Migration," Canadian Journal of Economics, 43(2), 663–682.

——— (2013a): "Barriers to Immigration and the Dynamics of Emigration," Journal of Macroeconomics, forthcoming.

——— (2013b): "Some Essentials of a Workable Guest-worker Program," International Economic Review, 54(2), 739–766.

Djajic, S., and M. S. Michael (2013): "Guest Worker Programs: A Theoretical Analysis of Welfare of the Host and Source Countries," The Journal of International Trade & Economic Development, 22(3), 454–475.

Djajic, S., M. S. Michael, and A. Vinogradova (2012): "Migration of Skilled Workers: Policy Interaction between Host and Source Countries," Journal of Public Economics, 96(11-12), 1015–1024.

Djajic, S., and R. Milbourne (1988): "A General Equilibrium Model of Guest-worker Migration: The Source-country Perspective," Journal of International Economics, 25(3-4), 335–351.

Docquier, F., B. L. Lowell, and A. Marfouk (2008): "International Migration by Educational Attainment with Gender Breakdown (1990 - 2000) - Release 2.1.," URL: http://perso.uclouvain.be/frederic.docquier/oxlight.htm [visited on 03.11.2011].

Docquier, F., and A. Marfouk (2006): "International Migration by Educational Attainment (1990-2000) - Release 1.1," in International Migration, Remittances and Development, ed. by C. Özden, and M. Schiff, pp. 151–199. Palgrave Macmillan, New York.

Docquier, F., and H. Rapoport (2012): "Globalization, Brain Drain, and Development," Journal of Economic Literature, 50(3), 681–730.

Docquier, F., C. Özden, C. Parson, and E. Artuc (2012): "A Global Assessment of Human Capital Mobility: The Role of non-OECD Destinations," Ires - institut de recherches economiques et sociales discussion paper no. 2012022.

Dostie, B., and P. Léger (2009): "Self-selection in Migration and Returns to Unobservables," Journal of Population Economics, 22(4), 1005–1024.

Dustmann, C. (2001): "Chapter 11 - Why go Back?: Return Motives of Migrant Workers," in International Migration: Trends, Policies, and Economic Impact, ed. by S. Djajic, vol. 3 of Routledge Contemporary Economic Policy Issues, pp. 229–249. Routledge.

Dustmann, C., I. Fadlon, and Y. Weiss (2011): "Return Migration, Human Capital Accumulation and the Brain Drain," Journal of Development Economics, 95(1), 58–67.

Dustmann, C., and A. Glitz (2011): "Chapter 4 - Migration and Education," in Handbook of The Economics of Education, ed. by S. M. Eric A. Hanushek, and L. Woessmann, vol. 4 of Handbook of the Economics of Education, pp. 327–439. Elsevier.

Dustmann, C., A. Glitz, and U. Schönberg (2011): "Referral-based Job Search Networks," IZA Discussion Papers 5777, Institute for the Study of Labor (IZA).

Dustmann, C., and O. Kirchkamp (2002): "The Optimal Migration Duration and Activity Choice after Re-migration," Journal of Development Economics, 67(2), 351 – 372.

Dustmann, C., and U. Schönberg (2012): "What Makes Firm-based Vocational Training Schemes Successful? The Role of Commitment," American Economic Journal: Applied Economics, 4(2), 36–61.

DUSTMANN, C., AND Y. WEISS (2007): "Return Migration: Theory and Empirical Evidence from the UK," British Journal of Industrial Relations, 45(2), 236–256.

EBENSTEIN, A. Y., A. E. HARRISON, M. MCMILLAN, AND S. PHILLIPS (2013): "Estimating the Impact of Trade and Offshoring on American Workers Using the Current Population Surveys," Review of Economics and Statistics, forthcoming.

ECKEL, C., AND H. EGGER (2009): "Wage Bargaining and Multinational Firms," Journal of International Economics, 77(2), 206–214.

EGGER, H., P. EGGER, AND U. KREICKEMEIER (2013): "Trade, Wages, and Profits," European Economic Review, 64(0), 332 – 350.

EGGER, H., AND U. KREICKEMEIER (2008): "International Fragmentation: Boon or Bane for Domestic Employment?," European Economic Review, 52(1), 116–132.

——— (2009): "Firm Heterogeneity and the Labor Market Effects of Trade Liberalization," International Economic Review, 50(1), 187–216.

——— (2012): "Fairness, Trade, and Inequality," Journal of International Economics, 86(2), 184–196.

EGGER, H., U. KREICKEMEIER, AND J. WRONA (2013a): "Offshoring Domestic Jobs," CESifo Working Paper Series 4083.

——— (2013b): "Offshoring Domestic Jobs," University of Tübingen Working Papers in Economics and Finance 50.

ENGELMANN, D., AND M. STROBEL (2004): "Inequality Aversion, Efficiency, and Maximin Preferences in Simple Distribution Experiments," American Economic Review, 94(4), 857–869.

ETHIER, W. J. (1982): "National and International Returns to Scale in the Modern Theory of International Trade," American Economic Review, 72(3), 389–405.

——— (1985): "International Trade and Labor Migration," The American Economic Review, 75(4), 691–707.

Eurofound (2012): "Fifth European Working Conditions Survey," Publications Office of the European Union, Luxembourg.

Feenstra, R. C. (2010): Offshoring in the Global Economy: Microeconomic Structure and Macroeconomic Implications. MIT Press, Cambridge, Mass. [u.a.].

Feenstra, R. C., and G. H. Hanson (1996a): "Foreign Investment, Outsourcing, and Relative Wages," in Political Economy of Trade Policy: Essays in Honor of Jagdish Bhagwati, ed. by R. C. Feenstra, G. M. Grossman, and D. A. Irwin, pp. 89–127. MIT Press, Cambridge/MA.

Feenstra, R. C., and G. H. Hanson (1996b): "Globalization, Outsourcing, and Wage Inequality," American Economic Review, 86(2), 240–45.

Feenstra, R. C., and G. H. Hanson (1999): "The Impact Of Outsourcing And High-Technology Capital On Wages: Estimates For The United States, 1979-1990," The Quarterly Journal of Economics, 114(3), 907–940.

Fehr, E., and A. Falk (1999): "Wage Rigidity in a Competitive Incomplete Contract Market," Journal of Political Economy, 107(1), 106–134.

Fehr, E., L. Goette, and C. Zehnder (2009): "A Behavioral Account of the Labor Market: The Role of Fairness Concerns," Annual Review of Economics, 1(1), 355–384.

Fehr, E., and K. M. Schmidt (1999): "A Theory Of Fairness, Competition, And Cooperation," The Quarterly Journal of Economics, 114(3), 817–868.

Felbermayr, G., V. Grossmann, and W. Kohler (2012): "Migration, International Trade and Capital Formation: Cause or Effect?," IZA Discussion Paper 6975.

Ferrer-i Carbonell, A. (2005): "Income and Well-being: An Empirical Analysis of the Comparison Income Effect," Journal of Public Economics, 89(5-6), 997–1019.

Gabaix, X., and A. Landier (2008): “Why Has CEO Pay Increased so Much?,” Quarterly Journal of Economics, 123(1), 49–100.

Gallaway, L. E. (1969): “Age and Labor Mobility Patterns,” Southern Economic Journal, 36(2), 171–180.

Galor, O. (1986): “Time Preference and International Labor Migration,” Journal of Economic Theory, 38(1), 1–20.

Geishecker, I., and H. Görg (2008): “Winners and Losers: A Micro-Level Analysis of International Outsourcing and Wages,” Canadian Journal of Economics, 41(1), 243–270.

Geishecker, I., M. Riedl, and P. Frijters (2012): “Offshoring and Job Loss Fears: An Econometric Analysis of Individual Perceptions,” Labour Economics, 19(5), 738–747.

Gersbach, H., and A. Schmutzler (2007): “Does Globalization Create Superstars?,” CEPR Discussion Paper 6222.

Giannetti, M. (2001): “Skill Complementarities and Migration Decisions,” LABOUR, 15(1), 1–31.

Gibson, J., and D. McKenzie (2011): “The Microeconomic Determinants of Emigration and Return Migration of the Best and Brightest: Evidence from the Pacific,” Journal of Development Economics, 95(1), 18–29.

——— (2012): “The Economic Consequences of ‘Brain Drain’ of the Best and Brightest: Microeconomic Evidence from Five Countries,” The Economic Journal, 122(560), 339–375.

Glitz, A. (2012): “Ethnic Segregation in Germany,” IZA Discussion Paper 6841.

Goos, M., A. Manning, and A. Salomons (2009): “Job Polarization in Europe,” American Economic Review P&P, 99(2), 58–63.

Görlitz, K., and J. Stiebale (2011): “The Impact of Product Market Competition on Employers Training Investments. Evidence from German Establishment Panel Data,” De Economist, 159(1), 1–23.

GOSS, E. P., AND C. PAUL (1986): "Age and Work Experience in the Decision to Migrate.," Journal of Human Resources, 21(3), 397–405.

GROIZARD, J. L., P. RANJAN, AND A. RODRIGUEZ-LOPEZ (2013): "Offshoring, Exporting, and Jobs," DEA Working Papers 61, Universitat de les Illes Balears, Departament d'Economia Aplicada.

GROSSMAN, G. M., AND E. HELPMAN (2008): "Fair Wages and Foreign Sourcing," in The Organization of Firms in a Global Economy, ed. by E. Helpman, D. Marin, and T. Verdier, pp. 273–310. Harvard University Press, Cambridge/MA.

GROSSMAN, G. M., AND E. ROSSI-HANSBERG (2008): "Trading Tasks: A Simple Theory of Offshoring," American Economic Review, 98(5), 1978–1997.

——— (2012): "Task Trade Between Similar Countries," Econometrica, 80(2), 593–629.

GRUBEL, H., AND P. J. G.LLOYD (1975): Intra-industry Trade : The Theory and Measurement of International Trade in Differentiated Products. Macmillan, London.

GRUBEL, H. B., AND A. D. SCOTT (1966a): "The International Flow of Human Capital," The American Economic Review, 56(1/2), pp. 268–274.

GRUBEL, H. G., AND A. D. SCOTT (1966b): "The Immigration of Scientists and Engineers to the United States, 1949-61," Journal of Political Economy, 74, 368–378.

HALL, A., AND M. TIEMANN (2006): "BIBB/BAuA Employment Survey of the Working Population on Qualification and Working Conditions in Germany 2006, suf 1.0," Discussion paper, Research Data Center at BIBB (ed.); GESIS Cologne, Germany (data access); Federal Institute of Vocational Education and Training, Bonn.

HANSON, G. H. (2009): "The Economic Consequences of the International Migration of Labor," Annual Review of Economics, 1(1), 179–208.

——— (2010): International Migration and the Developing Worldvol. 5 of Handbook of Development Economics, chap. 0, pp. 4363–4414. Elsevier.

Harms, P., O. Lorz, and D. Urban (2012): "Offshoring along the Production Chain," Canadian Journal of Economics, 45(1), 93–106.

Hellerstein, J. K., and D. Neumark (2008): "Workplace Segregation in the United States: Race, Ethnicity, and Skill," The Review of Economics and Statistics, 90(3), 459–477.

Helpman, E. (2006): "Trade, FDI, and the Organization of Firms," Journal of Economic Literature, 44(3), 589–630.

Helpman, E., and O. Itskhoki (2010): "Labour Market Rigidities, Trade and Unemployment," Review of Economic Studies, 77(3), 1100–1137.

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

Hendricks, L. (2001): "The Economic Performance of Immigrants: A Theory of Assortative Matching," International Economic Review, 42(2), 417–449.

Hogrefe, J., and J. Wrona (2013): "Trade, Tasks, and Training: The Effect of Offshoring on Individual Skill Upgrading," University of Tübingen Working Papers in Economics and Finance 64.

Hummels, D., J. Ishii, and K.-M. Yi (2001): "The Nature and Growth of Vertical Specialization in World Trade," Journal of International Economics, 54(1), 75–96.

Hummels, D., R. Jørgensen, J. R. Munch, and C. Xiang (2013): "The Wage Effects of Offshoring: Evidence from Danish Matched Worker-Firm Data," American Economic Review, forthcoming.

Hummels, D., J. Munch, L. Skipper, and C. Xiang (2012): "Offshoring, Transition, and Training: Evidence from Danish Matched Worker-Firm Data," American Economic Review P&P, 102(3), 424–428.

Iranzo, S., and G. Peri (2009): "Migration and Trade: Theory with an Application to the Eastern-Western European Integration," Journal of International Economics, 79(1), 1–19.

Johnson, K. M., P. R. Voss, R. B. Hammer, G. V. Fuguitt, and S. McNiven (2005): “Temporal and Spatial Variation in Age-Specific Net Migration in the United States,” Demography, 42(4), 791–812.

Jones, R. W. (2000): Globalization and the theory of input trade. MIT Press, Cambridge, Mass. [u.a.].

Jones, R. W., and H. Kierzkowski (1990): “The Role of Services in Production and International Trade: A Theoretical Framework,” in The Political Economy of International Trade: Essays in Honor of Robert E. Baldwin, ed. by R. W. Jones, and A. O. Krueger, pp. 31–48. Basil Blackwell, Cambridge/UK.

Karayalcin, C. (1994): “Temporary and Permanent Migration with and without an Immobile Factor,” Journal of Development Economics, 43(2), 197–215.

Kasahara, H., and B. Lapham (2013): “Productivity and the Decision to Import and Export: Theory and Evidence,” Journal of International Economics, 89(2), 297 – 316.

Katz, E., and O. Stark (1987): “International Migration under Asymmetric Information,” Economic Journal, 97(387), 718–26.

Katz, L. F., and D. H. Autor (1999): “Changes in the Wage Structure and Earnings Inequality,” in Handbook of Labor Economics, ed. by O. Ashenfelter, and D. Card, vol. 3 of Handbook of Labor Economics, chap. 26, pp. 1463–1555. Elsevier.

Katz, L. F., and L. H. Summers (1989): “Industry Rents: Evidence and Implications,” Brookings Papers on Economic Activity. Microeconomics, 1989, 209–290.

Keuschnigg, C., and E. Ribi (2009): “Outsourcing, Unemployment and Welfare Policy,” Journal of International Economics, 78(1), 168–176.

Kohler, W. (2003): “The Distributional Effects of International Fragmentation.,” German Economic Review, 4(1), 89.

Kohler, W. (2004a): “Aspects of International Fragmentation,” Review of International Economics, 12(5), 793–816.

——— (2004b): "International Outsourcing and Factor Prices with Multistage Production," Economic Journal, 114(494), C166–C185.

KREICKEMEIER, U., AND J. WRONA (2011a): "Two-way Migration Between Similar Countries," Discussion Papers 11/07, University of Nottingham, GEP.

——— (2011b): "Two-Way Migration between Similar Countries," University of Tübingen Working Papers in Economics and Finance 1.

KREMER, M. (1993): "The O-Ring Theory of Economic Development," The Quarterly Journal of Economics, 108(3), 551–575.

KRUEGER, A. B., AND L. H. SUMMERS (1988): "Efficiency Wages and the Inter-industry Wage Structure," Econometrica, 56(2), 259–293.

KRUGMAN, P. (1995): "Growing World Trade: Causes and Consequences," Brookings Papers on Economic Activity, 26(1, 25th A), 327–377.

KRUGMAN, P. R. (2000): "Technology, Trade and Factor Prices," Journal of International Economics, 50(1), 51–71.

KWOK, V., AND H. LELAND (1982): "An Economic Model of the Brain Drain," American Economic Review, 72(1), 91–100.

LAASER, C.-F., AND K. SCHRADER (2009): "Offshoring von Arbeitsplätzen: Ein empirischer Befund für Deutschland," Aussenwirtschaft: Schweizerische Zeitschrift für internationale Wirtschaftsbeziehungen, 64(2), 183–204.

LEAMER, E. (1995): "The Heckscher-Ohlin Model in Theory and Practice," Princeton Studies in International Economics 77, International Economics Section, Departement of Economics Princeton University,.

LEVY, F., AND R. J. MURNANE (2004): The New Division of Labor. Princeton University Press, Princeton/NJ.

Liu, R., and D. Trefler (2011): "A Sorted Tale of Globalization: White Collar Jobs and the Rise of Service Offshoring," NBER Working Paper 17559.

Loewenstein, G. F., L. Thompson, and M. H. Bazerman (1989): "Social Utility and Decision Making in Interpersonal Contexts.," Journal of Personality and Social Psychology, 57(3), 426 – 441.

Lowell, L. B. (2007): "Trends in International Migration: Flows and Stocks 1975-2005," Discussion paper, OECD Social, Employment and Migration Working Paper No. 58.

Lucas, R. E. J. (1978): "On the Size Distribution of Business Firms," Bell Journal of Economics, 9(2), 508–523.

Lucas, R. J. (1988): "On the Mechanics of Economic Development," Journal of Monetary Economics, 22(1), 3–42.

Luttmer, E. F. P. (2005): "Neighbors as Negatives: Relative Earnings and Well-Being," The Quarterly Journal of Economics, 120(3), 963–1002.

Martins, P. S., and Y. Yang (2013): "Globalised Labour Markets? International Rent Sharing across 47 Countries," British Journal of Industrial Relations, forthcoming.

Mattoo, A., I. C. Neagu, and A. Özden (2008): "Brain Waste? Educated Immigrants in the US Labor Market," Journal of Development Economics, 87(2), 255–269.

Matusz, S. J. (1996): "International Trade, the Division of Labor, and Unemployment," International Economic Review, 37(1), 71–84.

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

Melitz, M. J., and D. Trefler (2012): "Gains from Trade when Firms Matter," Journal of Economic Perspectives, 26(2), 91–118.

Méndez, F., and F. Sepúlveda (2012): "The Cyclicality of Skill Acquisition: Evidence from Panel Data," American Economic Journal: Macroeconomics, 4(3), 128–152.

Mesnard, A. (2004): "Temporary Migration and Capital Market Imperfections," Oxford Economic Papers, 56(2), 242–262.

Mitra, D., and P. Ranjan (2010): "Offshoring and Unemployment: The Role of Search Frictions Labor Mobility," Journal of International Economics, 81(2), 219–229.

Monarch, R., J. Park, and J. Sivadasan (2013): "Gains from Offshoring? Evidence from U.S. Microdata," Working Papers 13-20, Center for Economic Studies, U.S. Census Bureau.

Moser, C., D. M. Urban, and B. Weder di Mauro (2009): "Offshoring, Firm Performance and Establishment-Level Employment: Identifying Productivity and Downsizing Effects," CEPR Discussion Paper 7455.

Mountford, A. (1997): "Can a Brain Drain be Good for Growth in the Source Economy?," Journal of Development Economics, 53(2), 287–303.

Nekby, L. (2006): "The Emigration of Immigrants, Return vs. Onward Migration: Evidence from Sweden," Journal of Population Economics, 19(2), 197–226.

Ockenfels, A., and G. E. Bolton (2000): "ERC: A Theory of Equity, Reciprocity, and Competition," American Economic Review, 90(1), 166–193.

OECD (2007): Offshoring and Employment: Trends and Impacts. OECD, Paris.

OECD (2008a): "The Global Competition for Talent," OECD, Paris.

——— (2008b): "International Migration Outlook," OECD, Paris.

——— (2008c): "A Profile of Immigrant Populations in the 21st Century, Data from OECD Countries," OECD, Paris.

Ottaviano, G. I. P., and G. Peri (2012): "Rethinking the Effect of Immigration on Wages," Journal of the European Economic Association, 10(1), 152–197.

Ottaviano, G. I. P., G. Peri, and G. C. Wright (2013): "Immigration, Offshoring, and American Jobs," American Economic Review, 103(5), 1925–1959.

Ozden, C., C. R. Parsons, M. Schiff, and T. L. Walmsley (2011): "Where on Earth is Everybody? The Evolution of Global Bilateral Migration 1960-2000," World Bank Economic Review, 25(1), 12–56.

Paul, C. J. M., and M. Yasar (2009): "Outsourcing, Productivity, and Input Composition at the Plant Level," Canadian Journal of Economics, 42(2), 422–439.

Peri, G. (2012): "The Effect Of Immigration On Productivity: Evidence From U.S. States," The Review of Economics and Statistics, 94(1), 348–358.

Peri, G., and C. Sparber (2009): "Task Specialization, Immigration, and Wages," American Economic Journal: Applied Economics, 1(3), 135–69.

Pischke, J. (2001): "Continuous Training in Germany," Journal of Population Economics, 14(3), 523–548.

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

Rohrbach, D. (2009): "The BIBB/IAB-and BIBB/BAuA Surveys of the Working Population on Qualification and Working Conditions in Germany. Data and Methods Manual," Discussion paper.

Rooth, D.-O., and J. Saarela (2007): "Selection in Migration and Return Migration: Evidence from Micro Data," Economics Letters, 94(1), 90 – 95.

Rybczynski, T. M. (1955): "Factor Endowment and Relative Commodity Prices," Economica, 22(88), pp. 336–341.

Sattinger, M. (1975): "Comparative Advantage and the Distributions of Earnings and Abilities," Econometrica, 43(3), 455–468.

Schmitt, N., and A. Soubeyran (2006): "A Simple Model of Brain Circulation," Journal of International Economics, 69(2), 296–309.

Schmutzler, A., and H. Gersbach (2012): "Product Markets and Industry-Specific Training," RAND Journal of Economics, 43(3), 475 – 491.

Schwartz, A. (1976): "Migration, Age, and Education," Journal of Political Economy, 84(4), 701–720.

Shorrocks, A. F. (1980): "The Class of Additively Decomposable Inequality Measures," Econometrica, 48(3), 613–625.

Sjaastad, L. A. (1962): "The Costs and Returns of Human Migration," Journal of Political Economy, 70(5), 80–93.

Skaksen, J. R. (2004): "International Outsourcing when Labour Markets Are Unionized," Canadian Journal of Economics, 37(1), 78–94.

Spence, A. M. (1973): "Job Market Signaling," The Quarterly Journal of Economics, 87(3), 355–374.

Stark, O. (1995): "Return and Dynamics: The Path of Labor Migration When Workers Differ in Their Skills and Information is Asymmetric," Scandinavian Journal of Economics, 97(1), 55–71.

Stark, O., C. Helmenstein, and A. Prskawetz (1997): "A Brain Gain with a Brain Drain," Economics Letters, 55(2), 227–234.

——— (1998): "Human Capital Depletion, Human Capital Formation, and Migration: A Blessing or a 'Curse'?," Economics Letters, 60(3), 363 – 367.

Stark, O., and J. E. Taylor (1991): "Migration Incentives, Migration Types: The Role of Relative Deprivation," Economic Journal, 101(408), 1163–78.

Stolper, W. F., and P. A. Samuelson (1941): "Protection and Real Wages," The Review of Economic Studies, 9(1), pp. 58–73.

The Economist (2004): "Offshoring – The Rich World's Bangalore: How Britain Gains from Offshoring," November 11th.

——— (2009): "Idea – Idea Offshoring: Economists Argue That Offshoring Is a Win-Win Phenomenon," October 29th.

——— (2013): "Here, There and Everywhere," January 19th.

Tomiura, E. (2007): "Foreign Outsourcing, Exporting, and FDI: A Productivity Comparison at the Firm Level," Journal of International Economics, 72(1), 113–127.

Trax, M., S. Brunow, and J. Suedekum (2012): "Cultural Diversity and Plant-level Productivity," IZA Discussion Papers 6845.

von Greiff, J. (2009): "Displacement and Self-employment Entry," Labour Economics, 16(5), 556–565.

Wagner, J. (2012): "International Trade and Firm Performance: A Survey of Empirical Studies since 2006," Review of World Economics, 148(2), 235–267.

Wong, K.-y., and C. K. Yip (1999): "Education, Economic Growth, and Brain Drain," Journal of Economic Dynamics and Control, 23(5-6), 699–726.

Wrona, J. (2014): "Going Abroad for a Better Return," mimeo, University of Tübingen.

Xu, B. (2001): "Factor Bias, Sector Bias, and the Effects of Technical Progress on Relative Factor Prices," Journal of International Economics, 54(1), 5–25.

Yi, K. (2003): "Can Vertical Specialization Explain the Growth of World Trade?," Journal of Political Economy, 111(1), pp. 52–102.

Yitzhaki, S. (1979): "Relative Deprivation and the Gini Coefficient," The Quarterly Journal of Economics, 93(2), 321–24.