

Heterogeneous Workers, Trade, and Migration

Inga Heiland^a

Ifo Institute, Munich

Wilhelm Kohler^{b,c}

University of Tuebingen

August 5, 2016

(unfinished, please do not quote)

Abstract

We introduce horizontal skill differentiation among workers into a standard monopolistic competition model of trade. We show that with a non-convex technology this leads to monopsony power on the labor market as well as to endogenous average productivity through matching of workers to firms with different skill requirements. We assume a translog preferences and a "labor only" technology, and we focus on a symmetric equilibrium. Trade induces firm exit, thus aggravating the wage distortion from monopsony power on the labor market as well as lowering the average quality of matches between firms and workers. The gains from trade theorem survives, but welfare is non-monotonic in the level of real trade costs and trade increases wage inequality. Opening borders to international migration leads to two-way migration between similar countries. Migration leads to firm entry and an increase in the average quality of matches between firms, with an ambiguous effect on wage inequality. A "trade-cum migration" equilibrium is welfare-superior to a "free trade only" equilibrium, and welfare is monotonically increasing with lower real migration costs.

JEL codes: F12, F16, F22, J24

Keywords: two-way migration, gains from trade, heterogeneous workers

^aIfo Institute - Leibniz Institute for Economic Research at the University of Munich, Poschingerstrasse 5, 81679 Muenchen, Germany; heiland@ifo.de

^bUniversity of Tuebingen, Mohlstrasse 36, 72074 Tuebingen, Germany; Phone +49 (0)7071 2976013; wilhelm.kohler@uni-tuebingen.de. Corresponding author.

^cCESifo Munich, Germany, and GEP Nottingham, UK

1 Introduction

Workers are a heterogeneous lot. They have different innate abilities and choose different types and levels of skill formation. The outcome is both vertical and horizontal skill differentiation among workers. Economic well-being depends on good matches between workers' skills and their employment in different activities or firms. Obviously, the quality of these matches is determined by labor market institutions. Less obviously, it also depends on whether economies are connected to each other by trade and migration. In this paper, we analyze how horizontal skill heterogeneity among workers impacts on the welfare and inequality effects of trade and migration.

We develop a general equilibrium model of trade and migration where workers differ in terms of their specific skills. By skills, we mean an individual's relative ability to perform different types of tasks. By specificity, we mean that a certain pattern of such task-related abilities will be particularly useful for a certain product (or industry), and less so for others. Thus, for any given product there is a unique ideal pattern of task-related abilities, which we simply call the *ideal skill type*. The more a worker's specific abilities deviate from this ideal skill type, the lower her productivity when employed by a firm producing this product.¹ We focus on horizontal skill heterogeneity in assuming that all workers have the same absolute level of skills, but feature different relative abilities, i.e., different skill types. Moreover, we assume that both, the absolute level of skills and the degree of heterogeneity in skill types across workers is given exogenously, whether innately or from past educational investment.

We highlight two problems deriving from such worker heterogeneity. First, if the number of different skill types is large and if production involves a fixed cost, then the economy faces a *matching* problem. There will not be enough entry of firms for each worker to find for which her idiosyncratic skill type proves to be an ideal match. Part of the work force

¹ These characteristics of worker productivity are consistent with the concept of two-sided heterogeneity in the labor market where a multi-dimensional set of worker-specific skills needs to be matched with multi-dimensional skill requirements of jobs. Such environments are, for example, considered in Mandelbrot (1962); Rosen (1978); Moscarini (2001); Lazear (2009), and Lindenlaub (2016). Gathmann and Schönberg (2010) provide empirical evidence for the relevancy of skill-specificity and skill portability depending on a measure of distance between skill requirements of jobs.

will thus be employed in non-ideal matches, and the average quality of matches becomes an endogenous variable. And secondly, firms in such an economy will have *monopsony power* on the labor market. The reason is that for any expansion of employment a firm must hire workers with skills that are ever less suitable to its own product, and ever more suitable to the products of other firms. Attracting such workers requires offering a higher wage. Thus, firms face upward-sloping labor supplies and will therefore set wages below the marginal value product of labor.

In this paper, we show that these two problems, matching and monopsony power on the labor market, have profound implications for the normative effects of both, international trade and migration. To do so, we develop a model that combines horizontal skill-heterogeneity with the standard ingredients of monopolistic competition models of trade first identified by Krugman (1979): the variety effect, the pro-competitive effect, and the scale effect of international trade. However, in assuming symmetric firms, we deliberately rule out the selection-based productivity effect of trade emphasized by Melitz (2003). Instead, we highlight a productivity effect deriving from the average quality of matches between product characteristics and worker skills. Standard models of monopolistic competition emphasize a positive externality of firm entry due to love of variety, and a negative externality due to “business stealing” from incumbent firms. Our model highlights two additional distortions that derive from skill-specificity on the labor market: a positive externality of firm entry on the average quality of firm-worker-matches, and firms’ monopsony power on the labor market.

We employ a circular representation of continuous skill heterogeneity as in Amiti and Pissarides (2005).² Our analysis of demand and welfare is based on a translog expenditure function, which implies love of variety and endogenous markups.³ Labor is the only input

² Our model may also be seen in the tradition of spatial competition models, such as the circular city model developed by Vickrey (1964), Vickrey et al. (1999) and Salop (1979) or models of product differentiation in the spirit of Lancaster (1966). For similar trade applications, see Helpman (1981), Grossman and Helpman (2005) Eckel (2009a,b).

³ For the purpose of this paper, the terms product variety and product (or industry) may be used interchangeably. In some cases, our view of skill-specificity may fit product differentiation within an industry, so that the ideal skill relates to a certain product variety. In other cases, skill specificity may occur on a more aggregate level, which implies that ideal skills relate to a product or an industry. Thus, the notion of love of variety as implied by our translog expenditure function must flexibly be interpreted as related

in production, and all firms share the same technology, featuring a fixed cost and a constant marginal cost in terms of efficiency units of labor. Profit maximization and free entry of firms determine the equilibrium number of firms as well as the positions of these firms along the skill circle. In our model, firms' decision making involves two stages. Stage one involves the entry decision, combined with choosing firm-specific positions on the skill circle, and in stage two firms engage in Bertrand competition by jointly setting goods prices and wage rates per efficiency unit of labor. Workers maximize their earnings by sorting themselves into employment in different firms, based on firm-specific wage offers and skill-requirements. We assume an infinite number of potential entrants with zero outside options, whence entry is governed by a zero profit condition.

We assume that firms have knowledge of the distribution of workers over the skill circle as well as the degree of skill heterogeneity but do not know the individual worker's specific skills. In turn, workers have full knowledge about their skill types and the productivity effect of their skill distance to all firms once these have positioned themselves on the skill circle. When considering entry or exit, firms take as given the observed average quality of matches between worker skills and product requirements, thus ignoring the positive (negative) effect of entry (exit) on the average quality of worker-firm-matches.⁴

We use this model to analyze the effects of international trade as well as international migration on aggregate welfare as well as on wage inequality. Trade is analyzed assuming segmented national markets, and it comes in the two forms recently emphasized by Mrázová and Neary (2014). The first is the extensive margin where the number of countries trading with each other is increased above 1, which is autarky. The second is the intensive margin where a reduction in the level of trade costs leads to more trade between a given number of countries. Migration may take place in the form of cross-border hiring of workers, or in the form of workers moving to a foreign country in order to work for a foreign firm.

Trade has novel effects relative to existing literature. First, trade-induced firm exit causes a loss in efficiency through a lower average matching quality in the labor mar-

to product varieties or to products.

⁴ Such an externality is also present in Helsley and Strange (1990).

ket. And secondly, firm exit reduces competition in the labor market, leading to higher markups between wages and the marginal productivity of workers. However, comparing free trade with autarky (extensive margin) we prove that with translog preferences the conventional pro-competitive and variety effects dominate these adverse effects on the labor market. Thus, the gains from trade theorem survives. But the intensive margin involves a non-monotonicity: When moving from a prohibitive level of real trade costs (autarky) to zero trade costs (free trade) aggregate welfare is rising (falling) for high (low) initial levels of trade costs, provided that the number of countries is not too large. Both types of trade liberalization lead to an increase in wage inequality.

Novel results also arise for migration. With borders open for migration, some workers in each country find foreign firms that are better suited to their skills, and firms in each country find foreign workers that are better suited for their products. There will thus be an incentive for two-way migration. Moreover, this type of migration is unambiguously gainful, because it increases the average quality of matches while at the same time lowering wage markups in all countries. We show that the gains from lower wage markups even arise if migration costs are prohibitively high, since the mere potential of migration affects a firm's perceived labor supply elasticity. Moreover, we prove that any trade-cum-migration equilibrium always delivers higher aggregate welfare than an equilibrium with free trade alone. In contrast to piecemeal integration of goods markets, piecemeal integration of labor markets is unambiguously welfare increasing for all countries. And finally, in contrast to trade liberalization, the effect of migration on wage inequality is ambiguous.

Our paper contributes to several strands of literature. The first is the literature on whether trade and migration are substitutes or complements. Trade models highlighting endowment-based comparative advantage imply that they are substitutes, but if trade is driven by other forces they may be complements, as first emphasized by Markusen (1983). Empirical evidence strongly favors the view that trade and migration are “non-substitutes”; see Felbermayr et al. (2015). Our model identifies a novel cause of strong complementarity between trade and migration. We identify an incentive for migration which is present absent trade but increases with trade, and we demonstrate that this type of migration has effects opposite to those of trade.

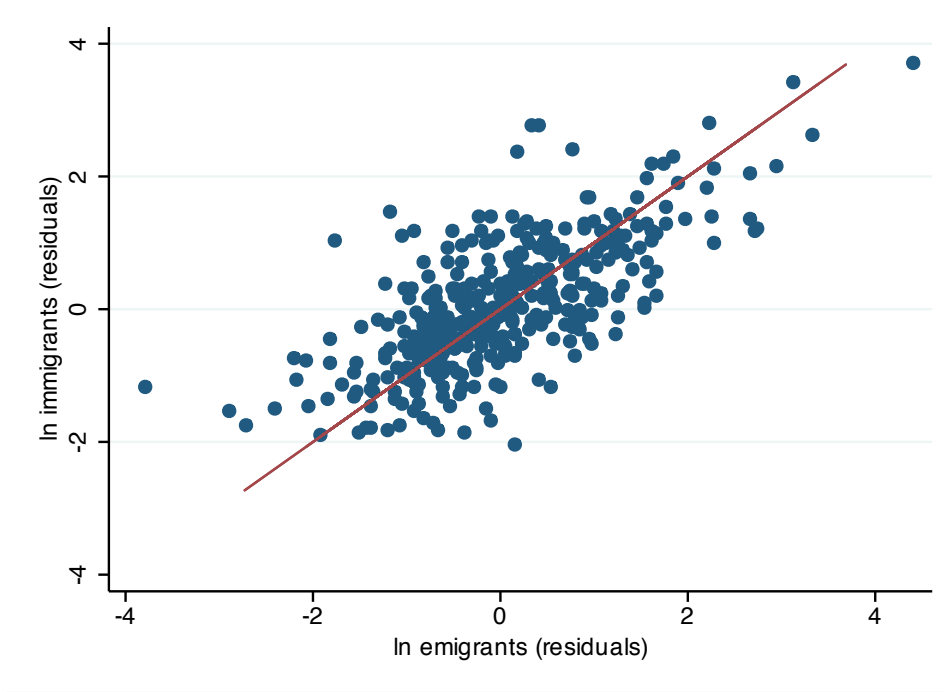


Figure 1: Bilateral stocks of immigrants and emigrants, OECD-DIOC 2000

A second strand of literature that we contribute to are the attempts to explain two-way migration between similar countries. Figure 1 demonstrates the empirical significance of this type of migration. It plots residuals from a gravity-type estimation of bilateral stocks (in logs) of emigrants and immigrants for all OECD-country pairs.⁵ Points clustering along the 45°-line reveal a marked two-way pattern of international migration among similar countries.⁶ This pattern of migration is difficult to explain. Schmitt and Soubeyran (2006) present a model predicting two-way migration of individuals within occupations.

⁵ The gravity equation is more difficult to rationalize for migration than for trade, but it is often used in the literature to explain cross-country patterns of migration, see. e.g. Ortega and Peri (2013), Mayda (2010), Beine et al. (2011). The data used in Figure 1 are from the *Database on Immigrants in OECD countries (DIOC)* provided by the OECD.

⁶ As a summary for the importance of two-way migration, Docquier and Marfouk (2006) report that 33.7% of the stock of skilled migrants residing in OECD countries stem from other high-income countries; see their Table 5.3. It is interesting to note that two-way migration was an important phenomenon also in the 19th century. It features prominently among the “Laws of Migration” that Ravenstein (1885) derives from a detailed account of 19th century migration flows between manufacturing districts in the UK. He calls it “counter-currents of migration,” and he insists that it is driven by “business considerations.” Our model may be interpreted as a theoretical account of such “business considerations.”

But in their model individuals with the same level of skills would never move in both directions, and even within occupations migration is observed only between countries that differ in their skill endowment. In Fan and Stark (2011), individuals suffer from social stigma arising from employment in an occupation of low social status, and humiliation from this stigma is felt to be less severe when working as a “foreigner” in the immigration country. This installs a two-way incentive for migration even between identical countries. Krickemeier and Wrona (2016) highlight vertical skill heterogeneity with a technology that requires formation of teams and features complementarity between skill levels of team members. If individual skill levels cannot be observed by firms and if migration is costly, then migration may serve as a signal for an above average level of skills. Because individuals in all countries gain from sending this signal, there will be two-way migration between identical countries. As in our model, migration alleviates a labor market imperfection, but it is driven by vertical skill differentiation, whereas we focus on horizontal skill differentiation. This allows us to investigate the link between migration and modern trade models featuring horizontal product differentiation.

As regards trade, our paper relates to a voluminous modern literature on gains from trade in the spirit of Krugman (1979) and Melitz (2003).⁷ Our contribution is to add horizontal skill heterogeneity among workers to an otherwise standard model of trade based on horizontal product differentiation, and to use this model to discuss the welfare and inequality effects of trade as well as migration, emphasizing novel effects deriving from endogenous wage markups as well as endogenous quality of worker-firm matching. To achieve these contributions, we simplify by assuming away firm-heterogeneity in productivity analyzed by Melitz (2003). We assume a translog expenditure function, which is nested in Arkolakis et al. (2015) and implies subconvex demand, as shown by Mrázová and Neary (2013). Thus, our model falls into the category of recent trade models delivering the familiar pro-competitive effects of trade.⁸

⁷ Arkolakis et al. (2012) has invigorated a renewed discussion of gains from trade. For a recent survey of this literature, see Costinot and Rodriguez-Clare (2014).

⁸ This categorization of trade models has recently emerged from attempts to move away from CES demand structures to allow for endogenous markups. Subconvexity of demand functions imply that any scenario leading to lower firm-sales in a certain market, such as entry of foreign firms into the domestic market, also leads to lower price markups on goods markets. For a detailed discussion, see Zhelobodko et al.

Finally, our paper also contributes to a recent strand of literature that explores the relationship between trade and matching in labor markets. Existing labor market literature has traditionally focused on vertical differentiation, emphasizing gains from positive assortative matching between firms and workers. Ohnsorge and Trefler (2007) as well as Costinot and Vogel (2010) discuss vertical skill differentiation that lead to perfect positive assortative matching in equilibrium, which implies that trade will not entail additional gains from better matching. In contrast to these papers, we discuss horizontal differentiation of workers with a variable average quality of matching. Focusing on a case where assortative matching is imperfect, Davidson et al. (2008) show that trade openness potentially enhances the degree of positive assortative matching. Using a large-scale Swedish data set, Davidson et al. (2012, 2014) demonstrate that this effect is empirically important. Our paper reinforces this point in showing that this channel also works with horizontal matching. Moreover, we incorporate these matching-based gains in a model that includes most of the other welfare channels highlighted by modern trade theory, and we use this model to analyze the matching effects not just of trade but also of migration.⁹

The remainder of the paper is organized as follows. In Section 2 we describe the general model framework and characterize the autarky equilibrium. In Section 3, we then discuss the effects of a transition from autarky to free trade and the scenario of piecemeal trade liberalization. In Section 4, we introduce labor mobility and analyze the effects of migration, first looking at a migration equilibrium and then comparing “trade cum migration” with trade alone. Section 5 concludes.

(2012) and Mrázová and Neary (2014, 2013).

⁹ Positive assortative matching also arises in Helpman et al. (2010) where firms are heterogeneous as in Melitz (2003). The reason is that firms may engage in costly screening to secure a minimum level of ability of hired workers, and more productive firms have a higher incentive to screen. But there is a crucial difference. In the present paper as well as in Davidson et al. (2008), skills are specific *ex ante*. In contrast, Helpman et al. (2010) assume *ex post* specificity: The ability of a worker revealed through firm-specific screening is specific to the match thus established. Ex ante, workers are identical in that each worker’s ability is drawn from the same distribution function at the time of hiring. Hence, a matching problem comparable to this paper does not exist in their model.

2 The Modeling Framework

Our model economy is endowed with a mass L of workers, which are differentiated by the *types* of skills they possess. A skill type is best thought of as a specific combination of abilities to perform different types of tasks. We assume that the entire space of skill types may be characterized by a circle with circumference $2H$, henceforth called the skill circle, whereby H measures the degree of horizontal skill differentiation present in the labor force. Each location on the circle represents a skill type, and types that are more similar are located closer to each other. This implies a continuous metric of similarity between different skill types. Moreover, using a circle to represent skill differences implies that each worker has the same average similarity to all other workers. Thus, skill heterogeneity is horizontal in nature. We assume a uniform distribution of the labor force L over the entire circle, which implies that a mass of $\frac{L}{2H}ds$ workers is located within an interval of length ds on the skill circle.

In order to set up production, a firm has to choose a certain location on the skill circle, which then determines that firm's ideal skill type. When working for this firm, workers will be differently productive, depending on the distance between their skills and the firm's ideal skill type. As will become evident below, this implies that firms are facing upward sloping supply schedules for efficiency units of labor. Thus, they have wage setting power, whereby the resulting markups depend on the precise pattern of firm locations on the skill circle.

Consumer preferences are described by a translog expenditure function, which implies love of variety. Firm behavior is structured in two stages. In *stage one*, potential entrepreneurs decide on whether to enter and, if so, where to locate on the skill circle. Setting up production at a certain point on the circle requires a fixed labor input α , defined in terms of efficiency units of the corresponding ideal skill type. In addition, production requires β units of this input per unit of the good produced. In *stage two*, firms set profit maximizing goods prices as well as wage rates, based on their market power on the goods as well as the labor market. We assume that firms pursue Bertrand strategies and that they are small enough to take aggregate variables as given. Stage two thus leads to a Bertrand-Nash equilibrium in prices and wage rates, conditional on the number

and skill positions of firms determined in stage one. Stage one decisions anticipate the Bertrand-Nash equilibria of stage two (subgame perfection). We assume free entry of an infinite number of potential entrepreneurs with zero outside options. Hence, equilibrium in stage one is determined by a zero profit condition.

The remainder of this section first looks at the second stage, including a detailed characterization of firms' wage setting power on the labor market. This is followed by a characterization of the zero profit equilibrium determined in stage one and by a characterization of an equilibrium where the economy is closed on both the goods and the labor market. A final subsection comments on the types of distortions involved in this equilibrium. Subsequent sections will then look at different scenarios of opening up the economy on the goods market and the labor market, respectively.

2.1 Price and wage setting with worker heterogeneity

2.1.1 Labor supply

For each variety that a firm may conceivably produce, technology is characterized by an ideal combination of skills that are needed to perform the required tasks. This combination corresponds to a unique point on the skill circle, to be chosen when a firm enters in stage one. Workers with skill types that deviate from this optimal type may still be employed to produce this variety, but will prove less productive. We model this through a function $f[d]$ which gives the number of efficiency units of labor delivered per physical unit of labor by a worker whose skills are represented by a point at distance d from the point representing the ideal type of skills. We assume that this function has the following properties: $f'[d] < 0$, $f'[0] = 0$, $f''[d] < 0$, and $f[d] = f[-d]$. This last property states that distance in either direction on the circle has the same effect. Efficiency units delivered by different types of workers are perfect substitutes. Without loss of generality, we set $f[0] = 1$.¹⁰

¹⁰One might ask why a firm should not always be able to secure the optimal combination of skills by employing convex combinations of workers embodying different combinations of skills. The answer is that doing so would entail a cost of communication between workers of different skill types. Therefore, other things equal, having the ideal skill type embodied in each worker is always less costly than combining different types of workers. The function $f[d]$ above may be interpreted as representing such cost of

We assume enforceable contracts between firms and workers, specifying the quantity of, and price for, efficiency units of labor. Each worker knows her skill distance from all firms positioned on the skill circle as well as the productivity schedule $f[d]$. Thus, she knows the income that she will earn per physical unit of labor when working for a certain firm offering a certain wage rate per efficiency unit. Each worker inelastically supplies one unit of physical labor. All workers sort themselves into employment by different firms so as to maximize their individual incomes, given firm-specific wage offers as well as their skill distance to these firms. For any pair of wage rates between two neighboring firms, there will thus be a marginal worker who is indifferent between the two firms as their wage offers amount to an equal income per physical unit of labor. All inframarginal workers earn wages above their outside options. This implies that the entire employment surplus is appropriated by workers, which is consistent with a zero profit equilibrium. The reason for why firms are unable to appropriate any employment surplus through wage discrimination is that they are unable to observe an individual worker's skill type. This effectively rules out paying each worker a wage rate equal to her outside option; this outside option is simply not known to the firm.

Figure 2 illustrates this type of worker sorting. It looks at a sector of the skill circle encompassing the location of three neighboring firms with optimal skill types s_i , s_{i+1} and s_{i+2} , which are at distances $2m_{i,i+1}$ and $2m_{i+1,i+2}$ from each other.¹¹ The concave curves depict firm-specific schedules $w_g f[d_g]$, $g = i, i+1, i+2$, giving the income that a worker at distance d_g from firm g 's position may expect to earn per physical unit of labor when working for this firm, given that it offers a wage rate per efficiency unit equal to w_g . The skill distance d_g is measured both to the left and the right from s_g . We refer to $w_g f[d_g]$ as firm g 's wage-income-schedule. Firm g 's wage rate is found as $w_g f[0]$. We define $d_{i,r}$ such that all workers in the interval $[s_i, d_{i,r}]$ prefer working for firm i to working for firms $i+1$ or $i+2$, and similarly for the interval $[s_i - d_{i,\ell}, s_i]$ to the left. In other words, $d_{i,r}$ and $d_{i,\ell}$ measure the skill distance between firm i 's ideal skill type and the marginal worker to the right and left, respectively, who is indifferent between working for firm i and its two

combining different skill types.

¹¹ In the symmetric equilibrium considered below, it will prove useful to use $m_{i,i+1}$ to indicate the half-distance (instead of the full distance) between firms i and $i+1$.

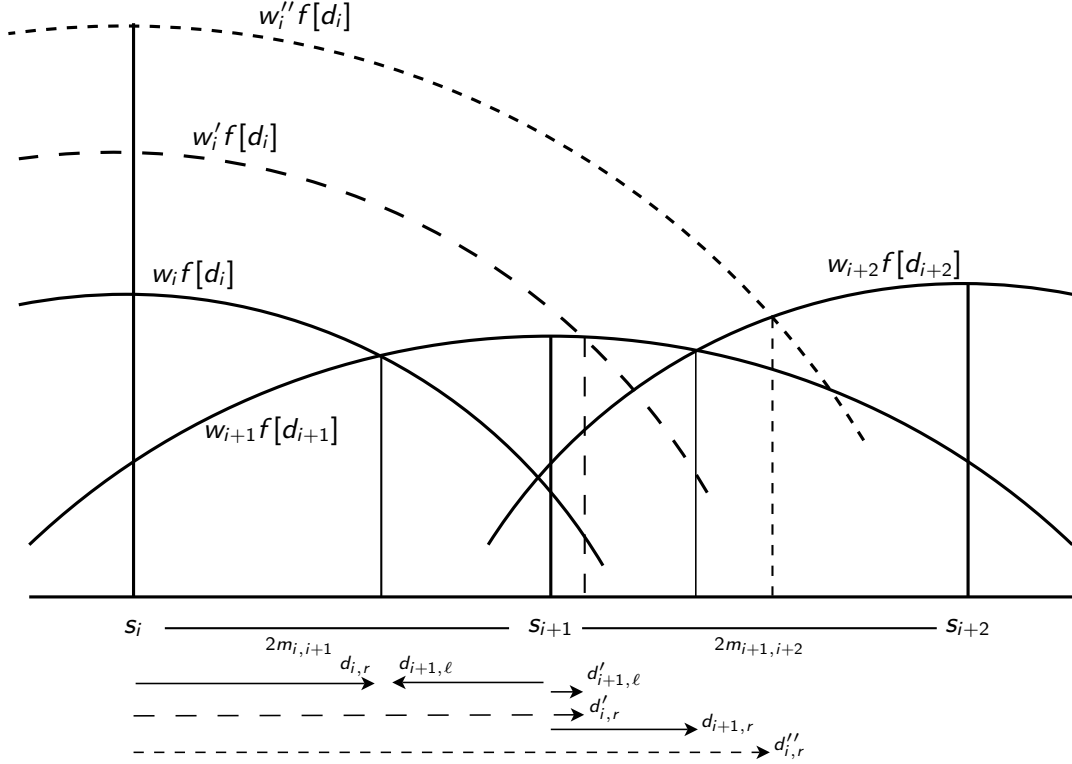


Figure 2: The wage schedules

neighboring firms. We shall also refer to this distance as the *skill reach* of firm i .

If firm i sets a wage equal to w_i , then its skill reach to the right is determined by the condition $w_if[d_{i,r}] = w_{i+1}f[2m_{i,i+1} - d_{i,r}]$; see the solid solid arrow at the bottom of Figure 2.¹² For a higher wage rate w'_i , a completely analogous condition $w'_if[d'_{i,r}] = w_{i+1}f[2m_{i,i+1} - d'_{i,r}]$ leads to a greater skill reach $d'_{i,r}$, indicated by the long-dashed arrow.¹³ With firm-specific wages w'_i and w_{i+1} firm i 's skill reach extends beyond s_{i+1} . Firm i would thus be able to hire firm $i + 1$'s ideal skill type, plus some workers to the right of s_{i+1} ,

¹²Note that we have replaced d_{i+1} by $2m_{i,i+1} - d_i$, in line with the aforementioned sorting of workers.

¹³Note that due to symmetry $f[2m_{i,i+1} - d'_i] = f[d'_i - 2m_{i,i+1}]$. This allows us to use the same condition determining marginal workers to the left and the right of s_{i+1} . Note also that workers at a distance $d'_{i,r} - 2m_{i,i+1}$ to the left of s_{i+1} would earn the marginal worker's income if working for firm $i + 1$, but they are better off working for firm i . The marginal worker is thus uniquely determined by the above condition.

up to point $s_i + d''_{i,r}$. In turn, firm $i + 1$ would be left employing workers in the interval $[s_i + d''_{i,r}, s_{i+1} + d''_{i+1,r}]$, with workers in the interval $[s_{i+1} + d_{i+1,r}, s_{i+2} + d_{i+2,r}]$ being employed by firm $i + 2$.¹⁴ Increasing its wage rate further to w''_i would allow firm i to out-compete firm $i + 1$ and start attracting workers from firm $i + 2$. The skill reach covered by w''_i is implicitly determined by $w''_i f[d''_{i,r}] = w_{i+2} f[2m_{i,i+1} + 2m_{i+1,i+2} - d''_{i,r}]$; see the short-dashed arrow in Figure 2. However, out-competing via high enough wages will never occur in the comparative static scenarios analyzed below, since any adjustment of the equilibrium number of firms occurs through firm exit and entry, driven by the condition of non-zero maximum profits.

In what follows we use \mathbf{w}_{-i} to denote the $N - 1$ vector of wage rates set by all firms other than i , such that the first element is the wage set by the first neighbor to its right, and so on until the N -th element, which is the wage set by the first neighbor to its left.¹⁵ Moreover, \mathbf{m}_i denotes the N -dimensional vector $(2m_{i-1,i}, 2m_{i,i+1}, \dots, 2m_{i-2,i-1})$, where we set $i - 1 = N$ if $i = 1$, and $i + 1 = 1$ if $i = N$. We shall henceforth refer to \mathbf{m}_i as the distance vector viewed at from firm i 's perspective. Consistently with the definition of \mathbf{w}_{-i} , $2m_{i-1,i}$ and $2m_{i,i+1}$ must be interpreted, respectively, as the distance between firm i and its first left-hand and first right-hand neighbor, and so on. It now follows from the above reasoning that the right-hand skill reach from firm i 's location on the skill circle may be written as $d_{i,r} = d_r[w_i, \mathbf{w}_{-i}, \mathbf{m}_i]$ and analogously for the skill reach to its left, $d_{i,\ell} = d_\ell[w_i, \mathbf{w}_{-i}, \mathbf{m}_i]$.¹⁶ Clearly, the skill reaches $d_{i,r}$ and $d_{i,\ell}$ are increasing in w_i and

¹⁴ The distance $d_{i+2,r}$ lies to the right of the range covered by Figure 2.

¹⁵ A well-defined labor supply function as derived in this subsection requires $N \geq 2$. We shall assume below that N is "large".

¹⁶ Note that the function $d_r[\cdot]$ is uniform across firms, but the value of this function, in general, will not. The function $d_r[w_i, \mathbf{w}_{-i}, \mathbf{m}_i]$ is implicitly defined as the solution to the following condition:

$$w_i f[d_{i,r}] = w_{i+j} f \left[\sum_{k=1}^j 2m_{i+k-1,i+k} - d_{i,r} \right] \quad \text{where} \quad (1)$$

$$j = \operatorname{argmax}_j \left\{ w_{i+j} f \left[\sum_{k=1}^j 2m_{i+k-1,i+k} - d_{i,r} \right] \mid 1 \leq j \leq N - 1 \right\},$$

where all m -terms are elements of the vector \mathbf{m}_i as defined above. This condition includes combinations of wage rates where firm i out-competes some of its nearest neighbors. In these expressions, $i + j$ indicates firm i 's relevant competitor employing the marginal worker at distance $d_{i,r}$ from firm i 's ideal type. The

weakly increasing in \mathbf{m}_i , but weakly decreasing in \mathbf{w}_{-i} .

The entire amount of efficiency units that firm i is able to attract through setting a wage rate w_i now follows as the integral over all efficiency units $f[d]$ from distance zero up to distance $d_{i,r}$ plus the corresponding integral from zero to $d_{i,\ell}$. Writing

$$\begin{aligned} L^{S,\ell} &= \int_0^{d_\ell[w_i, \mathbf{w}_{-i}, \mathbf{m}_i]} f[d] \frac{L}{2H} dd \\ L^{S,r} &= \int_0^{d_r[w_i, \mathbf{w}_{-i}, \mathbf{m}_i]} f[d] \frac{L}{2H} dd, \end{aligned}$$

firm i 's labor supply schedule now emerges as

$$L^S[w_i, \mathbf{w}_{-i}, \mathbf{m}_i] = \begin{cases} L^{S,\ell} + L^{S,r} & \text{if } d_{i,\ell} \leq -d_{i,r} \\ 0 & \text{else} \end{cases} \quad (2)$$

Intuitively, for a low enough wage rate w_i firm i 's labor supply will fall down to zero. This happens if the condition $d_{i,\ell} \leq -d_{i,r}$ is violated, in which case the skill reach covered by firm i 's neighbor to the left includes the skill reach that firm i is able to cover on its right. In other words, the overall distance covered by firm i then has zero measure. For wages above this threshold level, the firm faces a labor supply function for efficiency units which is increasing in its own wage. Moreover, the labor supply schedule is continuous in w_i except for points where a further increase in w_i reduces labor supply to the nearest competitor down to zero. But, as we have emphasized above, such out-competing of neighbors will never arise in the scenarios considered below. In what follows we shall use η_i to denote the elasticity of firm i 's labor supply function (2). Obviously, this elasticity is a function of w_i , \mathbf{w}_{-i} , and \mathbf{m}_i .

second line identifies the relevant competitor as the firm which is the first to meet firm i 's wage offer as the skill distance increases. Equivalently, it is the firm where a marginal (indifferent) worker is found at the shortest distance $d_{i,r}$ from firm i . A completely analogous determination holds for the left-hand skill reach $d_{i,\ell}$.

2.1.2 Goods demand

To complete our description of a firm's market environment, we next turn to goods demand. Individual k derives utility from consumption of a bundle $C[c_k]$ of differentiated varieties $c_k = [c_{1k}, \dots, c_{ik}, \dots, c_{Nk}]$, where N denotes the number of varieties available. Throughout the paper, we use brackets $[\cdot]$ to collect arguments of a function and parentheses to collect algebraic expressions. We assume that $C[c_k]$ is homogeneous of degree one, hence the logarithmic indirect utility function is given by

$$\ln V_k = \ln y_k - \ln P[p], \quad (3)$$

where $P[p] = P[p_1, \dots, p_i, \dots, p_N]$ is the minimum unit expenditure function for all varieties i , and y_k denotes income of individual k . Following Diewert (1974) and Bergin and Feenstra (2000), we assume that preferences are characterized by a symmetric translog expenditure function.¹⁷ The unit expenditure function is given by

$$\ln P[p] = \frac{1}{2\gamma N} + \frac{1}{N} \sum_{i=1}^N \ln p_i + \frac{\gamma}{2N} \sum_{i=1}^N \sum_{j=1}^N \ln p_i (\ln p_j - \ln p_i), \quad (4)$$

which is homogeneous of degree one. The parameter $\gamma > 0$ controls the degree of substitutability between varieties, a larger γ implying higher substitutability.¹⁸ Using Roy's identity, the Marshallian demand function for variety i can be derived as

$$q_{ik}[p, y_k] = \frac{\partial \ln P[p]}{\partial \ln p_i} \frac{y_k}{p_i} = \delta_i \frac{y_k}{p_i}, \quad (5)$$

¹⁷Recent applications of the symmetric translog expenditure system are Feenstra and Weinstein (2010), Arkolakis et al. (2010) and Rodríguez-López (2011). As Feenstra and Weinstein (2010) point out, another interesting feature of the this expenditure system is that it constitutes a second order Taylor approximation of any symmetric expenditure function.

¹⁸Feenstra (2003) derives a translog expenditure function of the type above allowing for a gap between the number of varieties *conceivably* available and the number of varieties available in a certain equilibrium. The specification used here, borrowed from Arkolakis et al. (2010), assumes that the number of varieties *conceivably* available is infinite and thus does not enter the expenditure function.

where

$$\delta_i = \frac{1}{N} + \gamma \left(\frac{1}{N} \sum_{j=1}^N \ln p_j - \ln p_i \right) \quad (6)$$

is the expenditure share for variety i . Thus, the preferences underlying the above expenditure function are homothetic. Inserting (6) into (5) and using Y to denote aggregate income, revenue from variety i , r_i , then follows as

$$r_i = \delta_i Y \quad \text{with} \quad (7)$$

$$\delta_i = \gamma \mathcal{W} \left[\exp \left\{ \frac{1}{\gamma N} + \overline{\ln p} \right\} \frac{q_i}{\gamma Y} \right], \quad (8)$$

where $\overline{\ln p} := \sum_i \ln p_i / N$ and $\mathcal{W}[\cdot]$ denotes the *Lambert function*.¹⁹ While (6) expresses the expenditure share as a function of $\ln p_i$, in (8) this share is expressed as a function of the quantity q_i ; Appendix A.1 has the details. Given our preferences, no two firms will produce the same variety, so that we may use i to indicate firms.

2.1.3 Pricing equilibrium

Armed with these representations of the firm's goods demand and labor supply, firm behavior in stage two may now be characterized by the following profit maximization problem:

$$\begin{aligned} & \max_{w_i} r_i - w_i L_i \\ \text{s.t.:} \quad & q_i = \frac{L_i - \alpha}{\beta} \quad \text{with } L_i = L^S[w_i, \mathbf{w}_{-i}, \mathbf{m}_i], \quad \text{and} \\ & q_i \geq 0. \end{aligned} \quad (9)$$

¹⁹The Lambert function $\mathcal{W}[z]$ defines the implicit solution to $xe^x = z$ for $z > 0$. Furthermore, it satisfies $\mathcal{W}_z = \frac{\mathcal{W}[z]}{(\mathcal{W}[z]+1)z} > 0$, $\mathcal{W}_{zz} < 0$, $\mathcal{W}[0] = 0$ and $\mathcal{W}[e] = 1$. Here as elsewhere in the paper, we use a subscript index to indicate partial derivatives whenever this proves convenient without causing confusion.

In (9), r_i must be seen as given in (7) and (8) above. The restriction ensures that firm i is on its labor supply function and produces a positive quantity. We proceed under the assumption that the non-negativity constraint is non-binding. The corresponding restrictions on the parameter space are discussed in Appendix (A.3). Note also that this maximization problem is conditional on the variables N and \mathbf{m}_i which are determined in stage one.

We assume that firms pursue Bertrand strategies on both the goods and the labor market, meaning that they take the prices and wages set by their competitors as given. Moreover, each firm is assumed to be small enough to take the average log price $\overline{\ln p}$ as well as aggregate income Y as given beyond its own influence. Under these assumptions the perceived price elasticity of demand for variety i emerges as

$$\varepsilon_i[p_i, \overline{\ln p}, N] := -\frac{d \ln q_i}{d \ln p_i} = 1 - \frac{d \ln \delta_i}{d \ln p_i} = 1 + \frac{\gamma}{\delta_i} > 0. \quad (10)$$

This elasticity depends on prices and the number of firms. Thus, the markup of prices over marginal cost, determined by stage two pricing decisions, will be an endogenous variable.

The first order condition for profit maximization requires that perceived marginal revenue is equal to perceived marginal cost. It follows that the optimal wage chosen by firm i observes

$$p_i = \frac{\varepsilon_i}{\varepsilon_i - 1} \frac{\eta_i + 1}{\eta_i} w_i \beta. \quad (11)$$

Pricing thus involves a double markup.²⁰ The first fraction in this pricing rule represents the markup that derives from the firm's price setting power on the goods market, and it is larger than 1 since $\varepsilon_i > 1$. From (10) and (6), we may write this markup as $\frac{\varepsilon_i}{\varepsilon_i - 1} = \left(1 + \frac{\delta_i}{\gamma}\right) = \mathcal{W} \left[\frac{\eta_i}{w_i(\eta_i + 1)} \exp \left\{ 1 + \frac{1}{\gamma N} + \overline{\ln p} \right\} \right]$, again using the Lambert function.²¹ The expression inside \mathcal{W} is a “summary measure” of the conditions that firm i faces on the labor market as well as the goods market. Given $\mathcal{W}_Z > 0$, a higher average log-price

²⁰ It is easy to verify that under the assumptions made, the second order condition is satisfied.

²¹ This derivation follows Arkolakis et al. (2010), for details see Appendix A.1.

of the firm's competitors and a lower degree of substitutability γ both lead to a higher markup. The same holds true for a smaller number of firms, whereas the markup is falling in perceived marginal cost. The second fraction in (11) represents the markup that derives from the firm's monopsony power on the labor market, where the firm faces a finite elasticity of supply $\eta_i < \infty$. Remember that η_i is a function of w_i , \mathbf{w}_{-i} and \mathbf{m}_i ; see (2) above.

Combining markup pricing as given in Equation (11) with firm-specific labor market clearing as given in the constraint in (9) implies $w_i = w[\mathbf{w}_{-i}, \mathbf{m}_i, N, \overline{\ln p}, Y]$. This is a best response function characterizing strategic interaction of firms in stage two.

When considering its optimal wage response, each firm takes as given the number of firms and the distance pattern determined in stage one and treats the macro-variables $\overline{\ln p}$ and Y as two constants. For ease of exposition, we shall drop the latter in what follows. The N best response functions in wages jointly determine N wage rates:

$$w_i^* = w^*[\mathbf{m}_i, N] \quad \text{for } i = 1, \dots, N \quad (12)$$

Given these wage rates, Equation (11) determines equilibrium prices $p_i^* = p^*[\mathbf{m}_i, N]$. Moreover, equilibrium wage rates determine equilibrium profits according to problem (9). We shall henceforth refer equilibrium profits as $\pi_i^* = \pi^*[\mathbf{m}_i, N]$. The conditions under which such an equilibrium exists and is unique may be summarized by the following lemma.

Lemma 1. *If marginal profits are positive for output levels in the neighborhood of zero, and if the profit function is quasiconcave in the wage rate, then there exists a unique equilibrium in the Bertrand game of wage and price setting in stage two. A sufficient condition for the profit function to be quasiconcave is that the labor supply function is concave. For any given labor market environment, quasiconcavity of profits obtains if the marginal cost β is sufficiently low.*

The proof of this lemma follows in Appendix A.2. The first condition rules out corner solutions in which some firms find it optimal, conditional on entry, not to produce at all. For a given labor market environment $(H, f[d], \mathbf{m}_i \text{ and } L)$, marginal profits are high if

the marginal cost β is low, and if the degree of substitutability in demand (captured by γ) is low. A low enough value of γ (low substitutability) ensures that the choke price for a new good is high enough, so that firms needing to set high prices due to tight labor market conditions $\mathbf{m}_i, \mathbf{w}_{-i}$ still face some demand for their good. Whatever the choke price, a low enough value of β will always ensure positive marginal profits. As regards quasiconcavity, β governs the weight of the curvature of the labor supply function in the curvature of the revenue function. Should labor supply be convex in the wage rate, a low enough value of β ensures that concavity of revenue in q_i leads to quasiconcavity of profits in w_i .

2.2 Entry decision and the equilibrium distance pattern

One way to think of the location choice in stage one is to view a firm's strategy space as a set of addresses it can choose, taking as given the addresses of other firms, and where the firm's pay-off is given by equilibrium profits as determined in the subsequent pricing game. This setup is chosen, for instance, in Economides (1989) and Vogel (2008) who analyze product differentiation, with consumers evenly distributed over a unit circle characterizing ideal product characteristics, and with utility quadratic and linear, respectively, in the distance between a consumer's and the firm's position on this circle. Analyzing existence and uniqueness of the equilibrium, one faces the challenge of characterizing the pricing game in stage two for alternative location choices in stage one, including asymmetric locations, that allows one to examine whether deviations from the symmetric location pattern are profitable. This is possible for linear or quadratic utility, respectively, in the setups considered by Economides (1989) and Vogel (2008).²² However, in our circular model of the labor market, where the labor supply function has a more general form, this proves

²²The contribution in Economides (1989) is to demonstrate the existence of a symmetric subgame-perfect equilibrium in a three stage game of (i) entry, (ii) choice of variety and (iii) pricing. Vogel (2008) develops a model similarly featuring three stages of decision making, but with firm heterogeneity in productivity and proving existence *and uniqueness* of the equilibrium. He demonstrates that, for linear utility, any one firm's market share and profit are determined only by its own productivity and by the average productivity across all firms. Moreover, he derives a relationship between the inter-firm pattern of productivities and the pattern of bilateral distances along the circle. Importantly, for homogeneous firms, this relationship implies a symmetric distance pattern.

intractable. We therefore choose a different approach, based on the observation that the address of a firm on the circle is not informative. We have shown above that equilibrium profits in the pricing game depend only on the distances to other firms, described by \mathbf{m}_i , and not by their positions as such. Knowing about this, a firm seems unlikely to consider alternative addresses for its own while assuming all other firms' addresses are held fixed. This would imply that the firm assumes it can influence the overall pattern of distances, which seems questionable. Indeed, complete symmetry of the circle implies that the only way for an individual firm to affect the distance pattern is by means of entry or exit. We therefore reduce the firm's choice of entry and the position on the skill circle to the decision to enter or not, given its beliefs about the type of distance patterns that it may rationally expect to face upon entry.

We assume that there is an infinite number \bar{N} of *potential* entrants. Given the circumference H , any given number of entrants N renders a set of infinitely many possible N -dimensional distance vectors \mathbf{m} between these N firms. In the following, we use \mathcal{M}^N to denote this set. It is an exhaustive description of possible labor market environments that a firm may face, if the number of entering firms is N . Now look at these distance vectors from the point of view of a certain firm i , using the above definition of \mathbf{m}_i as the N -dimensional vector $(2m_{i-1,i}, 2m_{i,i+1}, \dots, 2m_{i-2,i-1})$, where we set $i-1 = N$ if $i = 1$, and $i+1 = 1$ if $i = N$. Our approach to solving the entry game rests on the assumption that firm i views possible realizations of N distances as random variables, and that it forms beliefs about all conceivable distance patterns which we denote by $\mu_i[\mathbf{m}_i|N]$. The joint pdf μ_i assigns a unique probability to any conceivable distance vector (labor market environment) that firm i can possibly be confronted with upon entry, given that the total number of firms entering is N . A key observation now is that any one of these distance vectors implies a certain \mathbf{m}_{i+1} , \mathbf{m}_{i+2} , up to \mathbf{m}_{i-1} , suitably applying the above definition the firm-specific distance vector to firms $i+1$ up to $i-1$.

The strategy space for a firm is characterized by a binary decision variable \mathcal{I}_i , where $\mathcal{I}_i = 1$ indicates entry, and $\mathcal{I}_i = 0$ indicates non-entry. Firm i 's expected payoff, conditional on N is $\mathbb{E}_i[\pi^*[\mathbf{m}_i, N]]$, where \mathbb{E}_i denotes the expected value formed over all distance vectors viewed from firm i 's perspective and according to firm i 's set of conditional beliefs $\mu_i[\mathbf{m}_i|N]$. The key behavioral assumption now is that the firm's entry decision is based

on expected second-stage profits associated with all possible distance patterns, that is,

$$\mathcal{I}_i = \begin{cases} 1 & \text{if } \mathbb{E}_i[\pi^*[\mathbf{m}_i, N]] \geq 0 \text{ and } \nu_i[N] > 0 \\ 0 & \text{otherwise} \end{cases} \quad \text{for all } i = 1 \dots \bar{N} \quad (13)$$

In this decision rule, $\nu_i[N]$ denotes the probability of N according to firm i 's beliefs. Let $\mathcal{I}_i|N$ denote the decision outcome according to (13). Noting that $N = \sum_{j \neq i} \mathcal{I}_j + 1$, (13) is readily interpreted as a best response function. Considering the set of outcomes $\mathcal{I}_i|N$ for all possible values of N , an equilibrium is then described by

$$\sum_{i=1}^{\bar{N}} \mathcal{I}_i|N = N \quad (14)$$

It is important to note that the summation in (14) is from 1 to \bar{N} . Let ι be defined as an \bar{N} -dimensional vector $\{\mathcal{I}_1, \dots, \mathcal{I}_{\bar{N}}\}$, and let \mathfrak{I} denote the set of vectors ι that constitute an equilibrium according to (13) and (14). Then, each element of \mathfrak{I} implies a unique equilibrium number of entrants according to (14). However, any equilibrium number of entrants may be supported by more than one equilibrium vector ι .

To each and every $\mathcal{I}_i > 1$ appearing in some equilibrium vector ι corresponds a set of infinitely many labor market environments, i.e., distance vectors \mathbf{m}_i , with an associated set of beliefs $\mu_i[\mathbf{m}_i|N]$.²³ We now impose a consistency requirement on these beliefs that we argue is implied by the assumption of structural symmetry of firms, coupled with the assumption that firms are fully informed about the characteristics and logic of the circle (including that firms are symmetric). First, we argue that symmetry implies that beliefs are symmetric across firms, i.e., $\mu_i = \mu$ and $\nu_i = \nu$ for all i . Moreover, we argue that consistency requires that $\mu[\mathbf{m}_i|N]$ be zero for any distance vector that would imply negative profits for some of the N entrants. If some distance vectors are ruled out by this requirement, we re-scale probabilities of all remaining distance vectors, so that they sum up to 1. If for some N_0 all possible distance vectors $\mathbf{m}_i|N_0$ fall victim to this consistency requirement, firms know that if N_0 of them enter it will never be possible for all of

²³In the terminology of dynamic games with incomplete information, this set of possible labor market environments corresponds to an information set; see Mas-Colell et al., (1995, ch. 9).

them to make non-negative profits. Consistent beliefs about N then requires $\nu[N_0] = 0$, and the decision rule (13) implies that an outcome with entry of N_0 firms is ruled out. Moreover, we assume that firms internalize the logic of free entry and, therefore, attach zero probability to any number of firms N_0 , for which an equilibrium consistent with (13), (14), and $\mu[\mathbf{m}_i|N]$ exists for some $N > N_0$. We thus arrive at a set of *consistent beliefs*. Our refined definition of the entry equilibrium then reads

$$\max_N \left\{ \sum_{i=1}^{\bar{N}} \mathcal{I}_i | N = N \right\} \quad (15)$$

where optimal entry decisions $\mathcal{I}_i|N$ are based on *consistent beliefs*.

Given these requirements, we can prove under some restrictions on the parameter space discussed below that a symmetric equilibrium with $N^* > 0$ and zero profits, $\pi^*[\mathbf{m}_i, N^*] = 0$ is the only equilibrium for this entry game. Our proof, detailed in Appendix A.3, runs along the following logic. We first look at *symmetric* distance patterns. We demonstrate that there exists some N^* such that for the symmetric distance vector $\mathbf{m}_{\text{sym}}|N^*$ we have $\pi^*[\mathbf{m}_{\text{sym}}, N^*] = 0$ and for $N > N^*$ we have $\pi^*[\mathbf{m}_{\text{sym}}|N] < 0$. The logic of entry and equilibrium in (13) and (15), coupled with consistent beliefs, then implies that any (symmetric or asymmetric) outcome with $N < N^*$ cannot arise as an equilibrium. Next, turning to *asymmetric* distance patterns, we demonstrate that $N \geq N^*$ necessarily involves negative profits for at least one entrant. Therefore, by the exact same logic of (13) and (15) as well as consistency of beliefs, an asymmetric equilibrium with $N > N^*$ cannot arise as an equilibrium. Taken together, all of this implies that the symmetric distance pattern with $N = N^*$ is the only equilibrium.

A key element of this logic is the effect (on firms' maximum profits) of moving from a symmetric to an asymmetric distance pattern. This involves two channels. The first is a decrease in the average efficiency of labor use that comes with moving to asymmetry. Intuitively, starting out from a symmetric distance pattern, an asymmetric rearrangement of distances means that firms, on average, will see a reduction in labor supply at notionally unchanged wage rates. The second involves a change in labor supply elasticities that leads firms to adjust their wage markups. The disadvantage from the first channel of moving to

asymmetry is reinforced by the second if a firm's labor supply elasticity is falling, which implies a reduction in monopsony power. It turns out, however, that for an arbitrary change in the distance pattern, the effect on the labor supply elasticity is ambiguous. For the above logic to go through, we therefore need further conditions guaranteeing that any move from a symmetric to an asymmetric distance pattern will have a negative effect on at least one firm's maximum profits. We may summarize this by the following lemma.

Lemma 2. *Given that firms play entry strategies as described in (13), there exists a subgame-perfect equilibrium of entry as defined in (15) with a finite number of entering firms, symmetrically positioned on the skill circle, and this equilibrium is unique, provided that the following conditions are met: (i) Firms' beliefs about conceivable distance vectors and the number of entrants are consistent, (ii) the fixed cost are not too large relative to the size of the labor force and relative to the degree of product differentiation, so that at least one firm can charge a sufficiently high markup over marginal cost so as to break even, and (iii) the negative efficiency effect on the profits of a firm whose output decreases when moving from a symmetric to an asymmetric distance pattern is reinforced, or at least not undone, by the effect of this change on the firm's monopsony power on the labor market (labor supply elasticity).*

A detailed proof of this lemma is given in Appendix A.3. The general intuition for conditions (i), (ii), and (iii) has been discussed above. As detailed in Appendix A.3, condition (iii) is met, for instance, if the effect of moving from symmetry to asymmetry in the distance pattern on the slope of labor supply is sufficiently small in absolute value. Alternatively, it will be met if the marginal cost β is sufficiently low, meaning that the efficiency effect is large. Note that a low value of β is also required for Lemma 1. Condition (ii) ensures existence of a symmetric location pattern with positive second-stage equilibrium profits for at least one firm.

2.3 Autarky equilibrium

Having established symmetry of the equilibrium in stage one, we now turn to the determination of m , the equilibrium distance between any two representative firms, as well as

the solution for the goods price and the wage rate for the representative firm. In this section, we do this for the closed economy, thus paving the way for comparative analysis of various opening up scenarios in subsequent sections of the paper.

We first note that N , the number of firms is related to m through the circumference of the skill circle: $m = H/N$. Moreover, in a symmetric equilibrium we have $p_i = p$, with $\overline{\ln p} = \ln p$, as well as $w_i = w$. Next, we return to the above pricing rule (11), invoking symmetry in order to pin down the two markups on the goods and labor market. The elasticity of labor supply, given in (2) and evaluated at $w_i = w$, may be written as

$$\eta[m] := \left. \frac{\partial L_i^S}{\partial w_i} \frac{w_i}{L^S} \right|_{w_i=w} = -\frac{f[m]^2}{2F[m]f'[m]}, \quad (16)$$

where $F[m] := \int_0^m f[d]dd$. Our assumption that $f''[m] \leq 0$ ensures that the labor supply elasticity is falling in m .²⁴ Invoking symmetry in equation (6) simplifies the expressions for ε and δ , allowing us to write the profit maximizing price (11) as

$$p[m] = \rho[m]\psi[m]\beta, \quad (17)$$

$$\text{where } \rho[m] := 1 + \frac{1}{\gamma N[m]} \quad \text{and} \quad \psi[m] := \frac{\eta[m] + 1}{\eta[m]}. \quad (18)$$

In (17), we have normalized the wage per efficiency unit to 1.²⁵ Note that $\rho' > 0$ as well as $\psi' > 0$. Firms' monopsony power in the labor market increases as firms become larger and the number of firms falls. Equations (17) and (18) describe the first order condition on pricing: a higher distance between firms leads to a higher goods price p .

Next, we introduce θ to denote the average productivity of workers. Given a uniform

²⁴This follows directly from

$$\frac{\partial \eta[m]}{\partial m} = \frac{-f[m]}{F[m]} - \frac{f[m]^2}{2(F[m]f'[m])^2} (-f[m]f'[m] - F[m]f''[m]).$$

²⁵We are free to do so, since our equilibrium is homogeneous of degree zero in nominal prices. This can easily be seen from substituting (10) and (6) in (11), which yields $p_i = \left(1 + \frac{1}{\gamma N} + \overline{\ln p} - \ln p_i\right) \frac{\eta_i + 1}{\eta_i} w_i$.

distribution of the workforce around the circle, we have

$$\theta[m] = \frac{1}{m} \int_0^m f[d] dd. \quad (19)$$

Notice that we have $\theta_m = (f[m] - \theta[m])/m < 0$ since $f'[m] < 0$. Given our wage normalization, $\theta[m]$ represents average income per worker. In a situation where workers assume ex ante that each point on the skill circle has the same probability of representing an ideal skill-type of a firm, $\theta[m]$ may be interpreted as a worker's expected nominal income. Aggregate income emerges as $Y = L\theta[m]$, and output per firm is

$$q[m] = \frac{1}{N[m]} \frac{L\theta[m]}{p[m]}. \quad (20)$$

The zero profit condition requires

$$p[m] = \frac{\alpha + \beta q[m]}{q[m]}. \quad (21)$$

Without loss of generality, we may now scale units, such that $\beta = 1$. The labor market clearing condition may then be written as $\alpha + q[m] = \frac{L}{N[m]}\theta[m]$, and aggregate variable labor input is $N[m]q[m] = L\theta[m] - \alpha N[m]$. Substituting these expressions in (21), we obtain the following representation of the zero profit condition:

$$p[m] = g[m] := \frac{L\theta[m]}{L\theta[m] - \alpha N[m]}. \quad (22)$$

Note that $g[m] > 1$ is the usual measure of the degree of economies of scale, i.e., the ratio of average to marginal cost, applied to the economy at large. We have $g' < 0$.²⁶ With zero profits, this ratio must be equal to the price relative to marginal cost. With $w\beta = 1$ from our scaling and normalization, this is exactly what we have in equation (22). Intuitively, with a higher distance between firms, zero profit requires a lower price.

²⁶ Given our scaling assumption $\beta = 1$ and the normalization $w = 1$, $L\theta[m]$ is the economy-wide total cost, while $\alpha N[m]$ is the aggregate use of labor for fixed cost, both expressed in efficiency units of labor. Hence, the right-hand side of (22) is the aggregate equivalent to the ratio of average to marginal cost.

Combining the zero profit condition (22) with the Bertrand pricing equation in (17), we finally arrive at the following condition that determines m , the half-distance at which firms symmetrically locate on the skill circle in an autarky equilibrium:

$$g[m] = \rho[m]\psi[m]. \quad (23)$$

This is the core condition that we use in the subsequent comparative static analysis.

In all of the welfare results to be derived below, we take an ex ante view, assuming that workers regard each point on the circle as being equally likely to become an ideal type. Given a symmetric equilibrium, expected utility of a worker is then equal to

$$\ln V = \ln \theta[m] - \left(\frac{1}{2\gamma N[m]} + \ln p[m] \right). \quad (24)$$

Intuitively, this welfare measure is rising in income and the number of firms in the market, and is falling in the price of a typical variety of goods. Note, however, that all of these variables are depending on the equilibrium value of m . While we know from above that θ and N are both falling in m , the relationship between p and m is ambiguous at this stage of our analysis. As we shall see below, whether p rises or falls with m depends on the type of exogenous shock considered. Hence equation (24), while revealing, is no comparative static result. Before moving to a comparative static analysis in Section 3 below, we address the question of whether a laissez faire equilibrium incorporates an optimal value of m . Given the multiple distortions present in this economy, the expected answer is “No.” In the next subsection, we discuss these distortions in more detail, establishing the conclusion that the laissez faire equilibrium involves a sub-optimally large value of m , which implies excess firm entry.

2.4 Distortions

The equilibrium described above involves four distortions. (i) When considering market entry, firms fail to take into account the positive effect of their entry on welfare through a larger number of varieties. Following Dixit and Stiglitz (1977), this is often referred to as

“consumer-surplus distortion.” (ii) Moreover, potential entrants ignore the positive effect on average productivity arising from a better quality of matches in the labor market. This is novel in the present model, relative to standard models of monopolistic competition, and we call it the “productivity distortion.” Both, distortions (i) and (ii) constitute positive externalities, working towards insufficient entry in a laissez faire equilibrium. But entry also has negative externalities, having to do with markups on the goods and labor markets. More specifically, (iii) potential entrants anticipate both, a goods price markup as well as a wage markup, but fail to see that they will realize operating profits on such markups only at the expense of incumbent firms, due to the overall resource constraint. Following Mankiw and Whinston (1986), this may be called the “business-stealing” effect. And finally, (iv) potential entrants fail to anticipate that their entry will reduce the magnitudes of these same markups, due to enhanced competition. In a zero profit equilibrium, operating profits compensate for fixed cost, hence this “pro-competitive” effect, as well as the “business-stealing” effect, works towards excessive entry.

As is well known, in the standard CES version of the monopolistic competition model distortions (i) and (iii) offset each other and firm entry is efficient. In Appendix A.5 we show that in this model the net result of distortions (i)-(iv) is excess entry. Thus, the model inherits the “excess entry” result established by Salop (1979) for the circular city model.²⁷ Moreover the result is in line with Bilbiie et al. (2008), who find that in a monopolistic competition equilibrium with symmetric translog preferences the business-stealing effect dominates the consumer-surplus effect, giving rise to excess entry.²⁸ The excess-entry result plays a crucial role in the determination of the gains from globalization below, as those unfold partly through a mitigation of distortions.

²⁷ As an example for circular labor markets, see Helsley and Strange (1990).

²⁸ A further case in point has been established for preferences of the constant absolute risk aversion by Behrens and Murata (2012).

3 Symmetric trading equilibrium

In this section, we explore the gains as well as the wage inequality effects from trade. The first subsection compares autarky with free trade, whereby we introduce trade simply by allowing for the number of countries to increase beyond 1 (which is autarky) and allowing for firms in all countries to sell on all national markets without any border frictions. Mrázová and Neary (2014) call this the extensive margin of globalization. In the second subsection we then turn to the intensive margin of globalization by holding fixed (at 2) the number of countries, but allowing for trade to be costly and looking at marginal reduction of this cost. In both subsections, we rule out cross-border hiring of workers, i.e., ruling out international migration, which will be taken up in the next section. Both trade and migration is analyzed assuming countries to be fully symmetric, including the extent of worker heterogeneity, so as to clearly isolate the channels that emanate from horizontal worker heterogeneity as such.

Given worker heterogeneity, we must expect different workers to be affected differently in the trade and migration scenarios considered below. In order to address welfare effects, we must therefore specify the exact definition of aggregate welfare. We offer two alternative views, both leading to the same results. The first is to look at real income of the worker with average productivity $\theta[m]$, as given in (19). Given the assumed uniform distribution of workers over the skill circle, any increase in this income implies that workers whose income has fallen may be compensated through a (costless) lump-sum transfer system. In this sense, a rise in $\theta[m]$ may be interpreted as a Pareto improvement. The second definition of aggregate welfare exploits the fact that the model only determines m , the half-distance between two neighboring firms, but leaves the exact positioning of firms on the skill circle undetermined. It is therefore natural to treat the exact positioning of firms as unknown (ex ante) to workers, and to assume that all workers view all points on the circle as equally likely to become an ideal skill type of some firm. With these assumptions, the real income of the worker with average productivity $\theta[m]$ may be interpreted as a worker's expected utility. Since workers are assumed to be risk-neutral, an increase in expected utility will increase a worker's welfare.

3.1 Free trade

We assume that there are k symmetric countries and we denote the total number of firms worldwide by $N^T := kN$. Absent all barriers, prices for domestic and imported goods are equal, and given by

$$p[m] = \left(1 + \frac{1}{\gamma k N[m]}\right) \psi[m]. \quad (25)$$

This expression reflects the fact that firms now take into account foreign competitors, but it keeps the simplified form familiar from the autarky equilibrium; see (18). Absent all trade barriers, prices of imported and domestic varieties are fully symmetric, whence the price of any variety consumed is equal to the average price. In what follows, we define $\rho^T[m] := 1 + \frac{1}{k\gamma N[m]}$ as the goods price markup under free trade. It is obvious that $\rho^T[m] < \rho[m]$.

Total demand per variety remains unchanged, since the lower domestic demand is compensated by the larger number of countries:

$$q[m] = \frac{kL\theta[m]}{kN[m]p} = \frac{L\theta[m]}{N[m]p}. \quad (26)$$

The labor market clearing condition similarly remains unaffected. The equilibrium condition that determines m then follows as

$$g[m] = \rho^T[m]\psi[m]. \quad (27)$$

The following proposition summarizes the comparison between autarky, $k = 1$, and free trade among $k > 1$ countries.

Proposition 1. *Opening up to free trade among k symmetric countries (with $k > 1$) has the following effects, relative to an autarky equilibrium (where $k = 1$): (i) There is exit of firms in each country, with an increase in the total number of varieties available to the consumer. (ii) There is a higher wage markup, coupled with a lower price markup, but goods prices are unambiguously lower. (iii) Each country's labor market suffers from a*

fall in the average matching quality, with lower average income. (iv) Each country enjoys a higher real income and higher aggregate welfare. (v) Wage inequality increases.

Proof: A formal proof is relegated to Appendix A.6.1.

The increase in variety (i) and the pro-competitive effect on the goods market (ii) are standard results in trade models with monopolistic competition and endogenous markups. The novel insight here relates to adverse labor market effects: A lower number of domestic firms lowers the degree of competition on labor markets, increasing the wage markup. But the pro-competitive effect dominates, leading to lower prices under free trade than under autarky (ii). In addition, the exit of firms makes it more difficult for workers to find firms matching well with their skills, causing a reduction in the productivity of the average worker (iii). However, the variety and pro-competitive effects more than compensate for this negative productivity effect, making the economy better off under free trade than under autarky (iv). On account of $f'[m] < 0$ exit of some firms will reduce the lower bound of wages. Since the upper bound of wages is fixed at $f[0] = 1$, and given a uniform distribution of workers over the skill circle, this entails an increase in wage inequality.

This positive welfare effect in this proposition reflects the excess entry property of the laissez faire equilibrium, whence an exit of firms entails a first order welfare gain. This holds true whatever the cause of the exit. In the present scenario, this first order gain from $dm > 0$ is driven by opening up borders, $dk > 0$, which exerts a positive effect on household welfare through a larger number of product varieties available. However, workers are differently affected depending on their location on the skill circle. While the maximum wage rate paid to a worker remains unchanged, the ideal workers in the trade equilibrium are different from those of the autarky equilibrium. Statement (iv) of the proposition invokes the usual compensation argument in defining the aggregate welfare effect as the change in indirect utility of the worker who receives the average level of real income. Moreover, given the increase in m , the lower bound of wages paid will be falling. Hence, some workers even suffer a lower “nominal” wage rate because of a larger distance to the nearest firm on the skill circle.

3.2 Costly trade and piecemeal trade liberalization

The superiority of free trade to autarky does not imply that a piecemeal liberalization in a world with costly trade is always beneficial. We stick to the symmetric case, but for simplicity reduce the number of countries to $k = 2$, using an asterisk to denote the foreign country. Suppose that firms face iceberg transport cost $\tau > 1$ for exports. A domestic firm that sells q_i units on the domestic market and q_i^* units on the export market then needs a labor input equal to $\alpha + q_i + \tau q_i^*$.²⁹ We assume that markets are segmented, and therefore, firms can set market specific quantities independently. The firm thus maximizes profits with respect to the wage, which determines its labor supply and hence total output $\bar{q}_i = q_i + \tau q_i^*$, and with respect to the quantity sold on the domestic market observing $q_i^* = \frac{1}{\tau}(\bar{q}_i - q_i)$. Hence, it solves the following maximization problem:

$$\begin{aligned} \max_{w_i, q_i} \{ & r_i + r_i^* - w_i(\alpha + \bar{q}_i) \} \\ \text{s.t.:} \quad & r_i = \delta_i Y, \quad r_i^* = \delta_i^* Y^* \\ & \bar{q}_i = q_i + \tau q_i^* \quad \text{with} \quad q_i \geq 0 \text{ and } q_i^* \geq 0 \\ & \alpha + \bar{q}_i = L^S[w_i, \mathbf{w}_{-i}, \mathbf{m}_i] \end{aligned} \tag{28}$$

whereby

$$\delta_i = \frac{1}{N^T} + \gamma (\overline{\ln p} - \ln p_i) = \gamma \mathcal{W} \left[\exp \left\{ \frac{1}{\gamma N} + \overline{\ln p} \right\} \frac{q_i}{\gamma Y} \right] \quad \text{and} \tag{29}$$

$$\delta_i^* = \frac{1}{N^T} + \gamma (\overline{\ln p} - \ln p_i^*) = \gamma \mathcal{W} \left[\exp \left\{ \frac{1}{\gamma N} + \overline{\ln p} \right\} \frac{q_i^*}{\gamma Y} \right]. \tag{30}$$

In these equations, $\overline{\ln p} = \frac{1}{N} \sum_{j=1}^N \ln p_j + \frac{1}{N^*} \sum_{j^*=1}^{N^*} \ln p_{j^*}$ denotes the log average price of competitors, where j and j^* index firm i 's domestic and foreign competitors. Due to symmetry, the average log price is the same across markets.³⁰ The first order condition with

²⁹ Remember that we have scaled units such that the marginal production cost β equal unity.

³⁰ Due to symmetry, the expenditure functions are the same in both countries, but expenditure shares for domestic and imported goods are different. Expenditure shares are obtained by differentiation of the log expenditure function, i.e. $\delta_i := \frac{\partial \ln P}{\partial \ln p_i}$ and $\delta_i^* := \frac{\partial \ln P}{\partial \ln p_i^*}$, and then applying the same logic as outlined in Appendix A.1 to express them in terms of q_i and q_i^* , respectively.

respect to q_i commands that marginal revenue be equalized across markets, more specifically, that $p_i \left(\frac{\varepsilon-1}{\varepsilon} \right) = \frac{p_i^*}{\tau} \left(\frac{\varepsilon^*-1}{\varepsilon^*} \right)$. The first order condition with respect to w_i demands that, as above, marginal revenue equal perceived marginal cost.³¹ Acknowledging symmetric locations and identical wages and the normalization of the symmetric equilibrium wage to unity, we obtain the following optimal pricing conditions:

$$p = \frac{\varepsilon}{\varepsilon-1} \psi[m] \quad \text{with} \quad \frac{\varepsilon}{\varepsilon-1} = 1 + \frac{\delta}{\gamma} \quad (31)$$

$$p^* = \frac{\varepsilon^*}{\varepsilon^*-1} \psi[m] \tau \quad \text{with} \quad \frac{\varepsilon^*}{\varepsilon^*-1} = 1 + \frac{\delta^*}{\gamma} \quad (32)$$

The labor market clearing condition is

$$N[m] (\alpha + q[p, p^*, m] + \tau q^*[p, p^*, m]) = L\theta[m]. \quad (33)$$

In contrast to the autarky and the free trade case, the pricing conditions cannot be simplified further because individual firms' prices in (29) are not equal to average prices in the economy. The equilibrium skill reach of the representative firm, m , as well as domestic and export prices are determined by the system of equations (31), (32) and (33). This system is the analogue to the free trade equilibrium condition (27) above.³²

Our preferences imply that a finite level of real trade costs might be prohibitive. We denote this prohibitive level of trade costs by $\bar{\tau}$, and it is determined implicitly by $\delta_i^* = 0$ in (29). Note that with $\delta_i^* = 0$ the price elasticity of demand for foreign goods becomes infinite; see (10). Note also that high values of γ imply low values of $\bar{\tau}$. We may now state the following proposition on piecemeal trade liberalization.

Proposition 2. *For two identical countries in a trading equilibrium, a decrease in trade costs τ within the non-prohibitive range, $\tau \in [1, \bar{\tau})$, has the following effects: (i) There is exit of firms in each country. (ii) The price of imported varieties falls, but the change in the price of domestically produced goods is ambiguous: it falls at low initial levels of τ ,*

³¹ See Appendix A.6.2 for details.

³² As detailed in Appendix A.6.3, for comparative statics it proves convenient to rewrite the system of equations in terms of the endogenous variables m, W, W^* .

and it increases at high initial levels of τ . (iii) Aggregate welfare rises for sufficiently low initial levels of τ , and it falls for sufficiently high initial levels of τ . (iv) Wage inequality is increasing.

Proof: A formal proof is relegated to Appendix A.6.3.

Part (iii) of this proposition may seem puzzling at first sight. According to standard results on piecemeal trade liberalization, in this fully symmetric economy a *uniform* proportional reduction of trade barriers across all varieties should be a welfare increasing “liberalization formula”; see Fukushima (1979). The key difference here arises from the labor market distortion. Liberalization involves two opposing effects. First, a lower price for imported varieties leads firms to lower their price markup on domestic goods; a pro-competitive effect that is positive for welfare. Note that this effect arises even at the prohibitive margin with $\tau = \bar{\tau}$ where no imports take place in the initial equilibrium.³³ At the same time, however, as firms in both countries ship more output to foreign markets, they use up more resources for transport, which bids up wage rates and causes firm exits in both countries.³⁴ Fewer domestic firms imply larger markups on the labor market as well as a lower average quality of matches between firms and worker skills. The magnitude of this effect clearly depends on the initial level of trade costs; it is strongest at $\tau = \bar{\tau}$ and disappears for $\tau = 1$. The proof in the Appendix demonstrates that for $\tau = \bar{\tau}$ initially, the adverse labor market effect of a marginal reduction of τ dominates, not just in terms of higher prices for domestic varieties, but also in terms of welfare so that liberalization is welfare reducing. Since we can also demonstrate that free trade, $\tau = 1$, is better than autarky, $\tau = \bar{\tau}$, there is a threshold value $\tilde{\tau}$, with $1 < \tilde{\tau} < \bar{\tau}$, such that piecemeal liberalization starting from $\tau < \tilde{\tau}$ is unambiguously welfare increasing. Proposition 1 implies that $\tilde{\tau} > 1$.

³³ For $\tau = \bar{\tau}$ the trading equilibrium is quantitatively identical to the autarky equilibrium considered in Section 2.3. This can be shown by inserting the implicit solution for $\bar{\tau}$, obtained by setting $\delta^* = 0$, into the pricing condition (31). Yet, the disciplinary effect of a decrease in import prices works through $\ln p$ in equation (31), even if $\delta^* = 0$.

³⁴ This is the mechanism underlying the well-known home market effect for *asymmetric* countries, first noted by Krugman (1980). Of course, the home market effect as such does not arise here, since countries are assumed symmetric.

Invoking costless compensation, we use average income to evaluate aggregate welfare effects in an economy where heterogeneous workers are affected differently. Using the indirect utility function we see that welfare is affected by changes in prices of domestic and imported goods as well as by the change in m , which affects both average income and the number of available varieties. The change in welfare can be expressed as

$$\widehat{V} = \left(\frac{\partial \ln \theta}{\partial \ln m} - \frac{\partial \ln P}{\partial \ln m} \right) \widehat{m} - N\delta\widehat{p} - N\delta^*\widehat{p}^*. \quad (34)$$

The increase in m , induced by a decrease in trade costs, affects welfare negatively through a decrease in average income $\theta_m[m] < 0$ and a decrease in the number of firms, which implies an increase in the ideal price index $P_m[m] > 0$. The effects of changes in prices of domestic and imported goods are weighted by the respective expenditure shares. For a high initial level of trade costs the expenditure share for imported goods is small, so that consumers hardly benefit from the decrease in the price of imports, while being much affected by the change in the price of domestic goods, which is positive for a high initial level of trade costs. Hence, for a high initial level of τ the overall effect of a decrease in the trade cost level on welfare is negative. In contrast, for a low initial value of trade costs, the negative effect through a decrease in the number of firms becomes smaller, approaching zero as τ converges to one. Furthermore, the higher the expenditure share for imported goods, the more consumers benefit from lower import prices, the more important therefore also the competitive effect on domestic prices through the product market. Hence, we find a U-shaped relationship between welfare and the level of trade costs.

4 Migration

So far, we have analyzed the effects of product market integration under the assumption that workers are immobile across countries. Intuitively, if we allow for migration, then some workers in both countries will find their skills matching better with a firm in the foreign country. This constitutes an incentive for two-way migration. In this section, we show that a “free trade cum migration” equilibrium delivers higher welfare than a “free trade only” equilibrium. Furthermore, we show that under a slightly stronger assumption

about the curvature of $f[d]$, piecemeal integration of labor markets, unlike piecemeal trade liberalization, is unambiguously welfare increasing. Our model highlights two welfare increasing effects of migration: First, integration of labor markets reduces monopsony power, as domestic firms now compete for workers not only with other domestic firms, but also with foreign firms. Second, migration entails efficiency gains by improving the average matching quality, as workers in both countries are now able to find better skill matches for employment. Even though the number of available varieties might fall, compared to the free trade equilibrium, the efficiency gains and the pro-competitive gains on the labor market are always dominating, leading to a positive welfare effect.

4.1 Modeling migration

For simplicity, we consider the case of two symmetric countries, which implies the same number of firms in both countries, as well as equal prices and wages. This simplification allows us to focus on the part of migration that is related to the idea of skill mismatch. We deliberately ignore differences in average wages or in the cost of living that would clearly constitute migration incentives as well.³⁵ We model the cost of migration as reducing the productivity of a worker to a fraction $1 - \lambda$, if this worker moves to the other country. A domestic worker working for a domestic firm at distance d , delivering $f[d]$ efficiency units, thus delivers only $f[d](1 - \lambda)$ efficiency units when working for a foreign firm at the same skill distance d .³⁶ Symmetry of countries also means that the degree of skill differentiation among the labor force is the same.

We analyze international migration as an entry and pricing game with two stages and Bertrand behavior analogous to sections 2 and 3 above, occurring simultaneously in both countries, whereby each firm takes into account the possibility of hiring workers from the

³⁵ Labor mobility and free entry imply that there is also the possibility of an agglomeration equilibrium, where all workers and firms work and produce in the same country. This is ruled out, if the trade cost level is sufficiently low, compared to the cost of migration. Throughout this section, we assume free trade, hence our results are not impaired by instability. In the case of zero cost of both trade and migration, the equilibrium outcomes in the dispersed and the agglomeration equilibrium are the same in terms of prices and welfare.

³⁶ The proportionality assumption is convenient for modeling, yet it is not crucial. A general characterization of the specifications generating the results derived in this section is found in Appendix A.7.5.

other country. Conditions similar to the ones discussed in Section 2.2 ensure existence and uniqueness of the symmetric alternating location equilibrium for non-prohibitive migration cost; see Appendix A.7.1 for details. By alternating pattern, we mean any one firm facing two neighboring firms from the other country.

4.2 Labor supply with integrated labor markets

We continue using $2m$ to denote the skill distance between two firms located in the same country. In the alternating equilibrium the firm's direct competitor on the labor market, which is located in the foreign country, is then found at distance m in the skill space.³⁷ The sorting cut-offs, i.e., the maximum distances of native workers d^n and migrant workers d^m from their firms, are derived as follows. For a domestic firm i , taking the foreign wage w^* as given, the cut-off for native workers, $d_i^n = d_i^n[w_i, \bar{w}^*, m, \lambda]$, is determined by

$$w_i f[d_i^n] = \bar{w}^* f[m - d_i^n](1 - \lambda). \quad (35)$$

Analogously, the cut-off for migrant workers, $d_i^m = d_i^m[w_i, \bar{w}^*, m, \lambda]$, is determined by

$$w_i f[d_i^m] = \bar{w}^* f[m - d_i^m] \frac{1}{1 - \lambda}. \quad (36)$$

As the level of migration costs falls, the cut-offs converge. At $\lambda = 0$ they coincide at $m/2$.

Under symmetry, the employment and migration pattern will be as follows: The domestic firm employs domestic workers with skill-types in the interval $(\bar{s}_0 - d^n, \bar{s}_0 + d^n)$, and foreign workers (migrants) located in the interval $(\bar{s}_0 - d^m, \bar{s}_0 + d^m)$, while the foreign firm located at $\bar{s}_0 + m$ will employ foreign workers located in the interval $(\bar{s}_0 + m - d^n, \bar{s}_0 + m + d^n)$ and domestic workers (migrants) with skill types in the interval $(\bar{s}_0 + m - d^m, \bar{s}_0 + m + d^m)$. Notice that $d^n + d^m = m$.

³⁷Note that without migration $2m$ measures the distance to the nearest competitor.

The supply of efficiency units as a function of the firm's wage now emerges as

$$L^{ES,M}[w_i, \bar{w}^*, m, \lambda, L, H] = \frac{L}{H} \left(\int_0^{d_i^n[w_i, \bar{w}^*, m, \lambda]} f[d] dd + \int_0^{d_i^m[w_i, \bar{w}^*, m, \lambda]} f[d](1 - \lambda) dd \right) \quad (37)$$

where $d_i^n[w_i, \bar{w}^*, m, \lambda]$ and $d_i^m[w_i, \bar{w}^*, m, \lambda]$ are given by (35), (36), respectively. Let $d^n = d^n[m, \lambda]$ and $d^m = d^m[m, \lambda] := m - d^n[m, \lambda]$ denote the cut-offs in the symmetric equilibrium. These two variables measure the skill reach of a representative firm for domestic and foreign workers, respectively. With $\lambda > 0$, we have $d^m < d^n$. As before, m may be interpreted as a mismatch indicator, but the average distance between worker skills and a firm's ideal type across employment of domestic and foreign workers is now equal to $m/2$, whereas without migration it was equal to m . In a symmetric equilibrium, average productivity then emerges as

$$\theta^M[m, \lambda] := \frac{1}{m} \left(\int_0^{d^n} f[d] dd + \int_0^{d^m} f[d](1 - \lambda) dd \right). \quad (38)$$

By complete analogy to (16), the perceived elasticity of effective labor supply, evaluated at the symmetric equilibrium, can be derived as³⁸

$$\eta^M[m, \lambda] \big|_{w_i = \bar{w}^*} = \frac{2f[d^n]^2}{f'[d^n] + (1 - \lambda)f'[m - d^n]} \frac{-1}{\int_0^{d^n} f[d] dd + (1 - \lambda) \int_0^{d^m} f[d] dd}. \quad (39)$$

Note that the labor supply function is subject to the constraint $d^m[m, \lambda] \geq 0$, which ensures that both cut-offs lie in between the positions of the domestic and the foreign firm. This condition is equivalent to the condition that the migration cost level λ is not prohibitive.³⁹ As migration costs approach the prohibitive level, the supply of efficiency units of labor becomes equal to the supply under autarky. This is readily verified by inserting $d^n = m$ and $d^m = 0$ into (37).

³⁸ For details of the derivation see Appendix A.7.2

³⁹ Otherwise, if migration costs are too large relative to firm size, firms cannot attract any migrants in the first place and the supply curve looks different since they then compete again only with firms from the same country.

Interestingly, even if the level of migration costs is prohibitive, firm behavior is influenced by the mere potential of migration through the perceived elasticity of labor supply.⁴⁰ The possibility of attracting migrants by setting higher wages and thus increasing the supply of efficiency units implies that firms perceive a higher elasticity of supply, even if they do not employ any migrant in equilibrium. Let $\bar{\lambda}$ denote the prohibitive level of migration costs, determined by setting $d^m[m, \lambda] = 0$. The perceived wage elasticity of labor supply evaluated at $\bar{\lambda}$ is given by

$$\eta^M[m, \bar{\lambda}] = \frac{2f[m]^2}{f'[m] + (1 - \bar{\lambda})f'[0]} \frac{-1}{F[m]}. \quad (40)$$

Note that concavity of $f[d]$ is sufficient to ensure that $\eta^M[m, \bar{\lambda}]$ is larger than the elasticity of supply under autarky as given in (16). It is important to note at this point, that *with prohibitive* λ , the symmetric alternating location pattern does not constitute an equilibrium as defined in (15). The reason is precisely that in this situation, labor market competition is greater than in autarkic labor markets, while labor supply is the same. Firms can and will avoid this situation, for example, by simultaneously decreasing the distance to one of their foreign neighbors, leading to a de-facto autarkic labor market equilibrium. Nevertheless, the second-stage pricing equilibrium evaluated at the symmetric alternating location pattern and prohibitive cost of migration proves very helpful as a reference case for the equilibrium outcomes of the cases of non-prohibitive migration cost, for which the symmetric alternating location pattern does constitute the unique equilibrium of the two-stage game.

We show in Appendix (A.7.2) that $\eta^M[m, \lambda]$ increases as λ falls, provided that $f'''[d]$ is not too large. In what follows, we assume that this condition holds.⁴¹ By analogy to (18), we now use $\psi^M[m, \lambda] := (\eta^M[m, \lambda] + 1)/\eta^M[m, \lambda]$ to denote the wage distortion under

⁴⁰ We thank Vitor Trindade for pointing this out to us.

⁴¹ The reasoning behind this condition is as follows: A higher λ leads firms to increase the share of migrants employed by shifting d^n outwards and d^m inwards. If the curvature of $f[d]$ falls (in absolute terms) as the cut-offs move to the right, an increase in λ helps firms to avoid competition by employing more native workers in the range where the curvature of $f[d]$ is lower and fewer migrants in the range where the curvature of $f[d]$ is strong. We rule this out by assuming that the curvature does not decrease too much (in absolute terms) as the cut-off moves to the right.

migration. For a given level of m , the magnitude of this distortion is unambiguously lower with migration and $\lambda \in [0, \bar{\lambda}]$ than without.

In addition to the wage distortion, migration also affects the average quality of skill matches between workers and firms. It is obvious that for prohibitively high migration costs, $\lambda = \bar{\lambda}$, the average matching quality, as given in equation (38), is the same as under autarky, given in (19): $\theta^M[m, \bar{\lambda}] = \theta[m]$. Moreover, as we prove in the appendix, $\theta_\lambda^M < 0$. In other words, the matching quality unambiguously increases as λ falls, reaching $\theta^M[m, 0] = \theta[m/2]$ for frictionless migration where $\lambda = 0$. It is instructive to see how effective labor supply to a representative firm is affected by the cost of migration. Under frictionless migration, $\lambda = 0$, labor supply (37) emerges as

$$L^{ES,M} = 2 \frac{L}{H} \int_0^{\frac{m}{2}} f[d] dd = \frac{2L}{N^M} \theta^M[m, 0] = \frac{L}{N} \theta[m/2]. \quad (41)$$

Note that $N^M = \frac{2H}{m} = 2N$, where N is the number of firms in each country. Comparing this to the autarky case, both the number of firms and workers are doubled. However, we know from above that for $\lambda < \bar{\lambda}$ we have $\theta^M > \theta$. Hence, with migration firms face a larger supply of efficiency units of labor than under autarky. The reason is that, while employing the same number of workers as under autarky, each firm now finds workers with skills closer to its optimal type. Importantly, all of this is conditional upon a given level of m , which is determined by the firm entry condition. As we shall see below, equilibrium adjustment of the number of firms after opening up to migration, driven by a lower wage markup, might bring about firm exit which has a countervailing, negative effect on average productivity.

4.3 The “trade cum migration” equilibrium

We complete the description of a “trade cum migration” equilibrium for non-prohibitive migration costs by a zero profit condition that determines the number of firms or, equivalently, the distance between a neighboring domestic and foreign firm on the skill circle, which is now equal to m . We look at the case of free trade. As in (27), we formulate this condition as stating that the double markup is equal to the inverse of the degree of

economies of scale:

$$g^M[m, \lambda] = \rho^T[m] \psi^M[m, \lambda]. \quad (42)$$

In this equation, $\rho^T[m]$ denotes the free trade price markup over perceived marginal cost obtaining in a free trade equilibrium without migration. Under free trade, this markup simplifies to $1 + 1/(\gamma N^M)$, where N^M is the number of firms world-wide; see equation (25). Unlike the wage markup, the price markup is not affected by allowing for migration. The term $\psi^M[m, \lambda]$ denotes the wage markup in a migration equilibrium, as introduced above. The term $g^M[m, \lambda]$ on the left measures the degree of scale economies, taking into account the labor market clearing condition, which now reads as $\alpha + q = (m/H)L\theta^M[m, \lambda]$, as well as goods market clearing, which requires $q = L\theta^M[m, \lambda]/(pN)$. This measure thus reads as

$$g^M[m, \lambda] := \frac{L\theta^M[m, \lambda]}{L\theta^M[m, \lambda] - \alpha H/m}. \quad (43)$$

In order to understand the effects of labor market integration, we now proceed in two steps. We first look at a situation where migration is allowed in principle, but where the costs of migration are prohibitively large, $\lambda = \bar{\lambda}$, and compare this case with the equilibrium outcome under national labor markets. Importantly, as we have discussed above, the case of prohibitive migration cost we are referring to in this section describes the second-stage pricing equilibrium for a symmetric alternating location pattern, but does not constitute a subgame-perfect equilibrium as defined in (15). It serves only as a reference case for the second step, where we look at the effects of successively lowering the costs of migration, starting from non-prohibitive levels of λ , for which the symmetric alternating location pattern does constitute a subgame-perfect equilibrium.

Proposition 3. *Compared to a free trade equilibrium with national labor markets, welfare is unambiguously higher in the second-stage “trade cum migration” equilibrium with two symmetric countries, symmetric alternating firm locations, and a prohibitively high level of the cost of migration. The number of firms in each country is unambiguously smaller in both countries.*

Proof: The analytical details of the proof are relegated to Appendix A.7.3.

A key point to understand this proposition is that the excess entry property of the autarky equilibrium demonstrated in Section 2.4 is inherited by the second-stage zero-profit equilibrium with symmetric alternating location patterns for *any* $\lambda \in [0, \bar{\lambda}]$. While the productivity distortion is not affected as long as no one migrates, the wage markup is affected because firms perceive a larger elasticity of labor supply. By lowering the wage markup, opening up labor markets to migration implies that the number of firms in the zero-profit equilibrium is smaller, even if the cost of migration is prohibitively high. And given that the free trade equilibrium involves excessive firm entry, this entails a positive welfare effect. With a lower wage markup distortion relative to the productivity distortion, the allocation is now closer to the social optimum.

Proposition 4. *In a “trade cum migration” equilibrium with two symmetric countries, piecemeal integration of labor markets through a marginal reduction in the cost of migration has an ambiguous effect on the number of firms. However, it unambiguously leads to lower prices and an increase in welfare in both countries, irrespective of the initial level of migration costs $\lambda \in [0, \bar{\lambda}]$.*

Proof: The analytical details of the proof are relegated to Appendix A.7.4.

The intuition for this proposition is best grasped from Figure 3, which depicts the schedules $g^M[m, \lambda]$ and $\rho^T[m]\psi^M[m, \lambda]$, identifying the equilibrium value of m at the intersection, in line with the zero profit equilibrium condition (43). The vertical axis of Figure 3 may be interpreted as measuring goods prices. Remember that $g^M[m, \lambda]$ measures the inverse degree of scale economies, which is equivalent to the markup required for zero profits. An increase in m makes firms larger, but it also lowers the productivity of the average worker. The appendix shows that the size effect always dominates, whence the g^M -line is downward-sloping. The $\rho^T[m]\psi^M[m, \lambda]$ -line depicts the double markup, reflecting monopoly power on the goods market and monopsony power on the labor market, respectively. This line is unambiguously upward-sloping, as a lower number of firms (higher m) reduces both the perceived price elasticity of goods demand as well the perceived labor supply elasticity with respect to the wage rate. We know from

proposition 3 above that the intersection point for $\lambda = \bar{\lambda}$ involves a lower value of m than in the free trade equilibrium with national labor markets, which is determined by $g[m] = \rho^T[m]\psi[m]$.⁴²

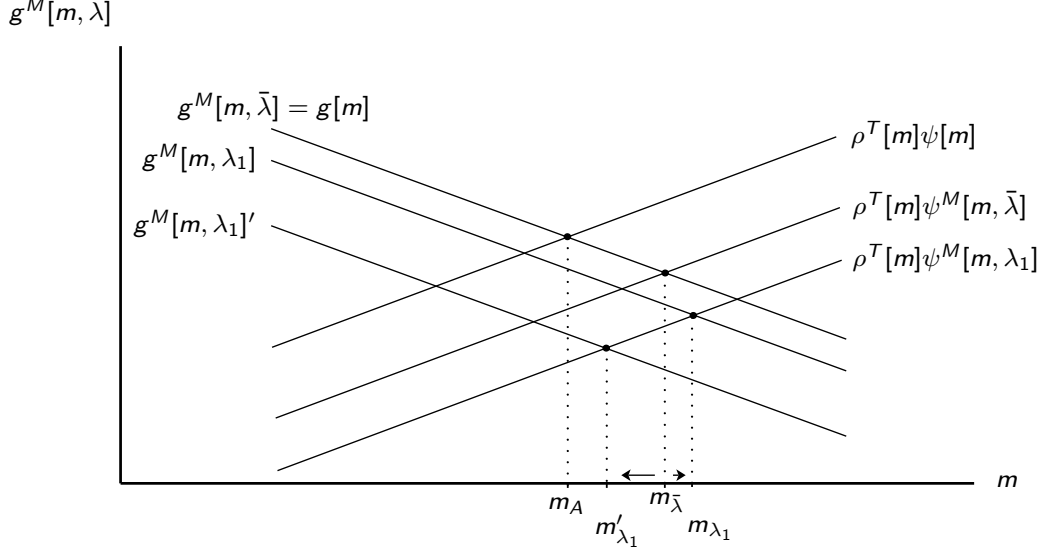


Figure 3: Comparative statics of the skill reach m

Now consider a reduction in λ from $\bar{\lambda}$ to $\lambda_1 \in [0, \bar{\lambda})$. For a notionally unchanged value of m , this improves the productivity of the average worker through a higher inframarginal surplus on migrant labor as well as through a resorting of workers from native employment into migration.⁴³ This means that the g^M -line is shifted down by a reduction in λ . As regards the markup schedule $\rho^T[m]\psi^M[m, \lambda]$, we have shown above that the perceived elasticity of labor supply increases with a lower cost of migration, meaning that for a notionally unchanged m firms charge a lower wage markup $\psi^M[m, \lambda]$. Thus, the markup schedule shifts down as well, rendering an ambiguous effect on m . In the figure, the case $g^M[m, \lambda_1]$ ($g^M[m, \lambda_1]'$) depicts a relatively weak (strong) shift in the g^M -line, leading to an increase (a decrease) in m . The ambiguity in the adjustment of m implies that wage inequality under migration can generally be lower or higher than in the free trade

⁴² Moving from an equilibrium with national labor markets to a “trade cum migration” equilibrium with $\lambda = \bar{\lambda}$ leaves g unaffected, $g[m] = g^M[m, \bar{\lambda}]$, while reducing the wage markup, $\psi[m] > \psi^M[m, \bar{\lambda}]$.

⁴³ Note that for a constant m the average skill distance between workers and their firm’s ideal type remains constant, equal to $m/2$. The productivity gain arises from savings in migration costs.

equilibrium. However, the equilibrium unambiguously moves down on the vertical axis, which implies lower goods prices.

The welfare effect is determined by the change in real income and the number of varieties. Real income is given by $\theta^M[m, \lambda]/p[m]$, where average “nominal” income is measured by $\theta^M[m, \lambda]$, the productivity of the average worker. Invoking the indirect utility function, the welfare effect of our scenario may be described as

$$\hat{V} = \frac{\partial \ln [\theta^M/p]}{\partial \lambda} \cdot d\lambda + \frac{\partial \ln [\theta^M/p]}{\partial m} \cdot dm - \frac{1}{4\gamma H} \cdot dm \quad (44)$$

The first term describes the direct effect of lower migration costs, $d\lambda < 0$, on real income. From the above we know that this term is unambiguously positive. The remaining terms involving dm are ambiguous in their entirety, because dm as caused by $d\lambda < 0$ is ambiguous. However, we know from the above discussion of the distortions present in this economy that the autarky equilibrium involves excess firm entry, and from the proof of proposition 3 we know that any second-stage “trade cum migration” equilibrium with symmetric alternating firm locations inherits this excess entry property. Therefore, the positive real income effect of firm exit in the second term must dominate the negative variety effect in the third term. In other words, if the equilibrium adjustment depicted in Figure 3 leads to $dm > 0$, then the overall effect of $d\lambda < 0$ on welfare is positive. If $dm < 0$, then the welfare effect is less straightforward. While the final term of this expression is then unambiguously positive, the first two terms seem ambiguous. However, we show in the appendix that the first two terms of (44) are unambiguously positive for any initial $\lambda \in [0, \bar{\lambda}]$, if we insert $dm = (\partial m/\partial \lambda) \cdot d\lambda$. Referring to our discussion subsequent to propositions 1 and 2, we repeat that individual households are affected differently, due to skill heterogeneity. Speaking of an aggregate welfare effect implies the existence of a costless (lump-sum) redistribution mechanism.

5 Conclusion

In this paper, we propose an important qualification to the common narrative of of variety-based gains from trade. Traditional models of monopolistic competition stress the impor-

tance of a large resource base for a large degree of product differentiation, if production is subject to a non-convex technology. By opening up to trade, even small countries may enjoy the benefits of a large resource base. Domestic firms may be driven out of the market, but this has no adverse effect. If anything, it increases the average productivity level through a positive selection effect.

This view neglects an important fact of modern manufacturing: Product differentiation relies on the availability of differentiated inputs, including non-traded inputs like labor. If producing a specific variety of a good requires a specific bundle of skills, then the skill-diversity of the labor force, rather than its size, determines the degree of product differentiation supplied by the market. In this paper, we have shown that trade is a somewhat less benign force in an environment where product differentiation is based on worker heterogeneity than portrayed in conventional models of monopolistic competition. In particular, trade-induced firm exit worsens the average quality of matches between the type of skills that workers bring to their firms and the specific skill requirements of the goods produced by these firms. In addition, product differentiation implies that firms have monopsony power in the labor market, whence trade-induced exit of firms increases the resulting distortion between the marginal productivity of labor and the wage rate. This works against the conventional pro-competitive effect of trade on the goods markets where trade lowers the markup between marginal cost and prices. Labor market integration gives rise to a migration incentive, whereby firms engage in cross-border hiring even under complete symmetry between countries. Migration essentially has effects that are opposite to those of trade.

We have developed a model which allows us to rigorously pin down these effects and to weigh them against the effects familiar from conventional models of monopolistic competition. In our model product differentiation is rooted in preferences represented by a translog expenditure function. When entering the market, firms decide upon which type of good to produce, based on a circular representation of skill heterogeneity among the work force, where each worker has the potential to serve as an “ideal” worker for a specific type of good. A non-convex technology implies a finite number of firms. A worker’s supply of efficiency units is inversely related to the distance between her skill-position on the circle and the ideal skill position of the firm she works for. Having positioned

themselves on the circle upon entry, firms engage in Bertrand competition on goods and labor markets, setting a double markup.

Using this model, we have explored both trade and migration scenarios. Comparing free trade with autarky in a symmetric many-country-world, we find that the variety and pro-competitive effects on goods markets unambiguously dominate the adverse effects from a lower average quality of matches between firms and workers and from higher markups on the labor market. Looking at piecemeal trade liberalization between two symmetric countries, we find an ambiguity: If liberalization takes place from a high initial level of trade costs, then it causes a lowering of aggregate welfare, whereas it increases aggregate welfare, if the initial level of trade costs is already below a certain threshold.

Starting from a free trade equilibrium in a symmetric two-country-world, integrating labor markets leads to two-way migration. Firms and workers in both countries face an incentive for cross-border hiring, even though the initial equilibrium features international wage equalization. Thus, our view of product differentiation based on worker heterogeneity generates a novel force of migration, contributing to an improved understanding of two-way migration, which looms large in the data but has so far lacked convincing explanation in standard models of migration. Interestingly, potential migration exerts a positive welfare effect on both countries, even if migration costs are prohibitively large. Contrary to piecemeal trade liberalization, a piecemeal reduction in the cost of migration is unambiguously welfare increasing. The reason is that it improves the quality of matches while at the same time lowering firms' monopsony power on labor markets. From the simple fact that trade and migration have opposite effects it also follows that trade and migration are complements, rather than substitutes. The model clearly advocates opening up labor markets simultaneously with trade liberalization.

Acknowledgements. This paper has greatly benefitted from helpful comments by Slobodan Djajic, Carsten Eckel, Gabriel Felbermayr, Gerhard Glomm, Gianmarco Ottaviano, Assaf Razin, Chris Parsons and Christian Schwarz. We also thank participants of the Trade and IO Seminar of the University of Munich, the 27th Annual Meeting of the EEA, the Midwest International Trade Conference Spring 2013, the 6th International Conference for Migration and Development, and the 6th Sinergia-Workshop (University of Zurich) on Economic Inequality and International Trade for valuable comments and suggestions. Wilhelm Kohler gratefully acknowledges financial support from the Volkswagen Foundation under the project "Europe's Global Linkages and the

Impact of the Financial Crisis: Policies for Sustainable Trade, Capital Flows, and Migration.”

References

- Amiti, M. and Pissarides, C. A. (2005). Trade and Industrial Location with Heterogeneous Labor. Journal of International Economics, 67(2):392–412.
- Arkolakis, C., Costinot, A., Donaldson, D., and Rodríguez-Clare, A. (2015). The elusive pro-competitive effects of trade. Working Paper 21370, National Bureau of Economic Research.
- Arkolakis, C., Costinot, A., and Rodríguez-Claire, A. (2010). Gains From Trade under Monopolistic Competition: A Simple Example with Translog Expenditure Functions and Pareto Distributions of Firm-Level Productivity. mimeo.
- Arkolakis, C., Costinot, A., and Rodríguez-Clare, A. (2012). New Trade Models, Same Old Gains? American Economic Review, 102 (1):94–130.
- Behrens, K. and Murata, Y. (2012). Trade, Competition, and Efficiency. Journal of International Economics, 87(1):1–17.
- Beine, M., Docquier, F., and Özden, C. (2011). Diasporas. Journal of Development Economics, 95(1):30–41.
- Bergin, P. R. and Feenstra, R. C. (2000). Staggered Price Setting and Endogenous Persistence. Journal of Monetary Economics, 45(3):657–680.
- Bilbiie, F. O., Ghironi, F., and Melitz, M. J. (2008). Monopoly Power and Endogenous Product Variety: Distortions and Remedies. NBER Working Paper 14383.
- Costinot, A. and Rodríguez-Clare, A. (2014). Trade theory with numbers: Quantifying the consequences of globalization. In Gopinath, G., Helpman, E., and Rogoff, K., editors, Handbook of International Economics, volume 4, chapter 4, pages 197–261.
- Costinot, A. and Vogel, J. (2010). Matching and inequality in the world economy. Journal of Political Economy, 118(4):pp. 747–786.
- Davidson, C., Heyman, F., Matusz, S., Sjöholm, F., and Zhu, S. C. (2014). Globalization and imperfect labor market sorting. Journal of International Economics, 94(2):177 – 194.
- Davidson, C., Heyman, F., Matusz, S., Sjöholm, F., and Zhu, S. C. (2012). Liberalized Trade and Worker-Firm Matching. American Economic Review, 102(3):429–34.
- Davidson, C., Matusz, S. J., and Shevchenko, A. (2008). Globalization and Firm Level Adjustment with Imperfect Labor Markets. Journal of International Economics, 75(2):295–309.
- Diewert, W. E. (1974). Applications of Duality Theory. In Intriligator, M. and Kendrick, D., editors, Frontiers of Quantitative Economics, volume 2, pages 106–171. North-Holland.

- Dixit, A. K. and Stiglitz, J. E. (1977). Monopolistic Competition and Optimum Product Diversity. American Economic Review, 67(3):297–308.
- Docquier, F. and Marfouk, A. (2006). International Migration by Educational Attainment (1990-2000) - Release 1.1. In Özden, C. and Schiff, M., editors, International Migration, Remittances and Development, volume 1, chapter 25. Palgrave Macmillan, New York.
- Eckel, C. (2009a). Endogenous Sunk Costs, Flexible Manufacturing and the Productivity Effects of International Trade. Scandinavian Journal of Economics, 111(2):369–386.
- Eckel, C. (2009b). International Trade, Flexible Manufacturing, and Outsourcing. Canadian Journal of Economics, 42(4):1449–1472.
- Economides, N. (1989). Symmetric Equilibrium Existence and Optimality in Differentiated Product Markets. Journal of Economic Theory, 47(1):178–194.
- Fan, C. S. and Stark, O. (2011). A theory of migration as response to economic stigma. International Economic Review, 52(2):549–571.
- Feenstra, R. C. (2003). A Homothetic Utility Function for Monopolistic Competition Models, without Constant Price Elasticity. Economics Letters, 78(1):79–86.
- Feenstra, R. C. and Weinstein, D. E. (2010). Globalization, Markups, and the U.S. Price Level. NBER Working Paper 15749.
- Felbermayr, G., Grossmann, V., and Kohler, W. (2015). Chapter 18 - migration, international trade, and capital formation: Cause or effect? In Chiswick, B. R. and Miller, P. W., editors, Handbook of the Economics of International Migration, volume 1, pages 913 – 1025. North-Holland.
- Fukushima, T. (1979). Tariff Structure, Nontraded Goods and Theory of Piecemeal Policy Recommendations. International Economic Review, 20(2):427–35.
- Gathmann, C. and Schönberg, U. (2010). How general is human capital? a taskbased approach. Journal of Labor Economics, 28(1):1–49.
- Grossman, G. M. and Helpman, E. (2005). Outsourcing in a Global Economy. Review of Economic Studies, 72(1):135–159.
- Helpman, E. (1981). International Trade in the Presence of Product Differentiation, Economies of Scale and Monopolistic Competition : A Chamberlin-Heckscher-Ohlin Approach. Journal of International Economics, 11(3):305–340.
- Helpman, E., Itskhoki, O., and Redding, S. (2010). Inequality and unemployment in a global economy. Econometrica, 78(4):1239–1283.
- Helsley, R. W. and Strange, W. C. (1990). Matching and Agglomeration Economies in a System of Cities. Regional Science and Urban Economics, 20(2):189–212.

- Kreickemeier, U. and Wrona, J. (2016). Two-way migration between similar countries. The World Economy, page forthcoming.
- Krugman, P. (1980). Scale Economies, Product Differentiation, and the Pattern of Trade. American Economic Review, 70(5):950–59.
- Krugman, P. R. (1979). Increasing Returns, Monopolistic Competition, and International Trade. Journal of International Economics, 9(4):469–479.
- Lancaster, K. J. (1966). A New Approach to Consumer Theory. The Journal of Political Economy, 74(2):pp. 132–157.
- Lazear, E. P. (2009). Firm-specific human capital: A skill-weights approach. Journal of Political Economy, 117(5):914–940.
- Lindenlaub, I. (2016). Sorting multi-dimensional types: Theory and application. unpublished manuscript.
- Mandelbrot, B. (1962). Paretian Distributions and Income Maximization. The Quarterly Journal of Economics, 76(1):pp. 57–85.
- Mankiw, N. G. and Whinston, M. D. (1986). Free Entry and Social Inefficiency. RAND Journal of Economics, 17(1):48–58.
- Markusen, J. R. (1983). Factor Movements and Commodity Trade as Complements. Journal of International Economics, 14(3-4):341–356.
- Mas-Colell, A., Whinston, M. D., and Green, J. R. (1995). Microeconomic Theory. Oxford University Press, New York, Oxford.
- Mayda, A. M. (2010). International Migration: A Panel Data Analysis of the Determinants of Bilateral Flows. Journal of Population Economics, 23(4):1249–1274.
- Melitz, M. J. (2003). The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity. Econometrica, 71(6):1695–1725.
- Moscarini, G. (2001). Excess worker reallocation. The Review of Economic Studies, 68(3):593–612.
- Mrázová, M. and Neary, J. P. (2013). Not so demanding: Preference structure, firm behavior, and welfare. University of Oxford, Department of Economics, Working Paper 691.
- Mrázová, M. and Neary, J. P. (2014). Together at last: Trade costs, demand structure, and welfare. American Economic Review, 104(5):298–303.
- Ohnsorge, F. and Trefler, D. (2007). Sorting It Out: International Trade and Protection with Heterogeneous Workers. Journal of Political Economy, 115(5):868–892.

- Ortega, F. and Peri, G. (2013). The effect of income and immigration policies on international migration. Migration Studies, 1(1):47–74.
- Ravenstein, E. (1885). The laws of migration. Technical Report 2.
- Ródriguez-López, J. A. (2011). Prices and Exchange Rates: A Theory of Disconnect. The Review of Economic Studies.
- Rosen, S. (1978). Substitution and division of labour. Economica, 45(179):235–250.
- Salop, S. C. (1979). Monopolistic Competition with Outside Goods. The Bell Journal of Economics, 10(1):pp. 141–156.
- Schmitt, N. and Soubeyran, A. (2006). A Simple Model of Brain Circulation. Journal of International Economics, 69(2):296–309.
- Vickrey, W. (1964). Microstatics. Harcourt, Brace and World, New York.
- Vickrey, W. S., Anderson, S. P., and Braid, R. M. (1999). Spatial Competition, Monopolistic Competition, and Optimum Product Diversity. International Journal of Industrial Organization, 17(7):953–963.
- Vives, X. (2001). Oligopoly Pricing: Old Ideas and New Tools, volume 1. The MIT Press, 1 edition.
- Vogel, J. (2008). Spatial competition with heterogeneous firms. Journal of Political Economy, 116(3):pp. 423–466.
- Zhelobodko, E., Kokovin, S., Parenti, M., and Thisse, J.-F. (2012). Monopolistic competition: Beyond the constant elasticity of substitution. Econometrica, 80(6):2765–2784.

Appendix

A.1 Expenditure share and markup

Starting out from Equation (6), we obtain

$$\delta_i = \frac{1}{N} + \gamma \overline{\ln p} - \gamma \ln \frac{\delta_i Y}{q_i}$$

by inserting $p_i = \frac{\delta_i Y}{q_i}$. This can be rewritten as

$$\frac{\delta_i}{\gamma} + \ln \frac{\delta_i}{\gamma} = \frac{1}{\gamma N} + \overline{\ln p} - \ln \frac{Y}{q_i} - \ln \gamma$$

and applying the Lambert function $\mathcal{W}[z]$ which is defined as the solution to $xe^x = z$ or, equivalently, to $\ln x + x = \ln z$, yields $\delta_i = \delta[q_i, \overline{\ln p}, N, Y]$ as defined in Equation (8).

Similar logic can be applied to obtain an explicit solution for the optimal price as determined by (11). Defining perceived marginal cost as $\tilde{w}_i := [(\eta_i + 1)/\eta_i]w_i\beta$ and observing (6) and (10) the first-order condition (11) can be written as

$$\frac{p_i}{\tilde{w}_i} + \ln p_i = 1 + \frac{1}{\gamma N} + \overline{\ln p}. \quad (\text{A.1})$$

The left-hand side is an implicit function of the profit maximizing price p_i . Rewriting (A.1) as

$$\frac{p_i}{\tilde{w}_i} + \ln p_i - \ln \tilde{w}_i = 1 + \frac{1}{\gamma N} + \overline{\ln p} - \ln \tilde{w}_i \quad (\text{A.2})$$

and applying the Lambert function to the left-hand side, we obtain the following explicit solution for p_i

$$p_i = \mathcal{W} \left[\tilde{w}_i^{-1} \exp \left\{ 1 + \frac{1}{\gamma N} + \overline{\ln p} \right\} \right] \tilde{w}_i. \quad (\text{A.3})$$

which implies that the price markup obeys

$$\frac{\varepsilon_i}{\varepsilon_i - 1} = \mathcal{W} \left[\frac{\eta_i}{w_i(\eta_i + 1)} \exp \left\{ 1 + \frac{1}{\gamma N} + \overline{\ln p} \right\} \right].$$

A.2 Existence and uniqueness of the pricing equilibrium (Lemma 1)

We invoke the Index Theory approach outlined in (c.p. Vives, 2001, p. 48) to proof that under certain restrictions on the parameter space, there is a unique solution to second stage game. It follows, that $\pi^*[\mathbf{m}, N]$ is unique. The Index Theorem approach is based on the Poincaré-Hopf Index Theorem, which implies that a solution to a system of reaction functions is unique if

- i). payoff functions are quasiconcave in firms' own strategies,
- ii). the strategy space is convex and compact and all equilibria are interior,
- iii). the Hessian is negative definite at the equilibrium point.

The proof proceeds as follows: We show that condition i) holds if the elasticity of marginal labor supply is not too large and condition ii) holds if marginal revenue is positive for output levels arbitrarily close to zero. Then, we show that condition iii) is always fulfilled in a transformed game where firms' strategies are *log* wages. Since conditions i) and ii) also hold in the transformed game, the Index Theorem implies that the transformed game has a unique solution. Since $\ln w$ is a positive monotone transformation of w for $w > 0$, this, in turn, implies uniqueness of the solution to the original game.

i) Quasi-concavity of the profit function. In the second stage, firms take the distance pattern \mathbf{m}_i , aggregate income Y and the average log price $\overline{\ln p}$ as given and determine their optimal wage as best response against other firms' wage choices \mathbf{w}_{-i} by maximizing profits as in (9). The set of permissible strategies is bounded from below by $w_{\alpha i} := L^{-1}[\alpha, \mathbf{w}_{-i}, \mathbf{m}_i] > 0$, which denotes the wage-level where the second constraint in (9) binds.⁴⁴ Moreover, firms never set wages above the *choke price* divided by the marginal labor requirement β , where

$$p_{choke} := p_{choke}[N] = \lim_{q_i \rightarrow 0} mr_i = \exp \left[\frac{1}{\gamma N} + \overline{\ln p} \right] \quad \text{with}$$

$$mr_i := mr[q_i, N, Y, \overline{\ln p}] = \frac{\partial r_i}{\partial q_i} = \frac{r_i}{(W_i^q + 1)q_i} > 0$$

⁴⁴ $w_{\alpha i}$ approaches zero as all firms lower their wages towards zero. We assume that $w_{\alpha i}$ is positive because otherwise $\ln w$ (which we will be working with below) is not defined. This assumption has no bearing on the equilibrium outcomes. Moreover, excluding the possibility of zero wages can be justified by assuming that at a zero wage, workers prefer not to work and hence, firms need to pay at least the reservation wage.

that is, the price at which demand approaches zero. Hence, there is a common upper bound on firms' strategies $\bar{w} = \frac{\exp[\frac{1}{\gamma N} + \ln p]}{\beta}$ and so firm i 's strategy set is $S_i = [w_{\alpha i}, \bar{w}]$.⁴⁵ In the above and the subsequent equations, we use $W_i^q := \mathcal{W} \left[\exp \left\{ \frac{1}{\gamma N} + \ln p \right\} \frac{q_i}{\gamma Y} \right]$ to denote the markup. Moreover, we henceforth denote with $L_i := L^S[w_i, \mathbf{w}_{-i}, \mathbf{m}_i]$ firm i 's labor supply.

A sufficient condition for $\pi[w_i, \mathbf{w}_{-i}, \mathbf{m}_i, N]$ to be quasiconcave in the firm's own strategy is that $\frac{\partial^2 \pi_i}{\partial w_i^2} < 0$ whenever $\frac{\partial \pi_i}{\partial w_i} \geq 0$. Marginal profits are given by

$$\frac{\partial \pi_i}{\partial w_i} = \left(\frac{mr_i}{\beta} - w_i \right) \frac{\partial L_i}{\partial w_i} - L_i \quad \text{if } \frac{\partial L_i}{\partial w_i} \geq 0,$$

where

$$\frac{\partial L_i}{\partial w_i} = \frac{L}{2H} \sum_{c=\ell, r} f[d_{i,c}] \frac{dd_{i,c}}{dw_i} > 0 \quad \text{with} \quad \frac{dd_{i,c}}{dw_i} = \frac{f[d_{i,c}]}{-w_i f'[d_{i,c}] - w_c f'[2m_{i,c} - d_{i,c}]} > 0. \quad (\text{A.4})$$

Marginal profits change in the firm's wage according to

$$\frac{\partial^2 \pi_i}{\partial w_i^2} = \frac{\partial mr_i}{\partial q_i} \frac{1}{\beta^2} \left(\frac{\partial L_i}{\partial w_i} \right)^2 - 2 \frac{\partial L_i}{\partial w_i} + \left(\frac{mr_i[q_i]}{\beta} - w_i \right) \frac{\partial^2 L_i}{\partial w_i^2} \quad \text{if } \frac{\partial L_i}{\partial w_i} \geq 0, \quad (\text{A.5})$$

where

$$\frac{\partial^2 L_i}{\partial w_i^2} = \frac{L}{2H} \left(\sum_{c=\ell, r} 3f'[d_{i,c}] \left(\frac{dd_{i,c}}{dw_i} \right)^2 + (w_i f''[d_{i,c}] - w_c f''[2m_{i,c} - d_{i,c}]) \left(\frac{dd_{i,c}}{dw_i} \right)^3 \right) \leq 0 \quad (\text{A.6})$$

$$\frac{\partial mr_i}{\partial q_i} = - \frac{mr_i}{q_i} \frac{W_i^q (W_i^q + 2)}{(W_i^q + 1)^2} \quad (\text{A.7})$$

Rewriting Equation (A.5), we obtain

$$\frac{\alpha + \beta q_i}{\beta q_i} \cdot |\varepsilon_{mr,i}| \geq \frac{\eta_{mls,i}}{\eta_i} \quad (\text{A.8})$$

⁴⁵ Note that p_{choke} is the upper bound on marginal revenue and $w\beta$ is the lower bound of firms' marginal cost. Hence, wages above p_{choke}/β entail negative marginal profits independent of other firms' choices and are thus strictly dominated.

as a sufficient condition for quasiconcavity, where

$$\varepsilon_{mr,i} = \frac{\partial mr_i}{\partial q_i} \frac{q_i}{mr_i} = -\frac{W_i^q(W_i^q + 2)}{(W_i^q + 1)^2} \leq 0$$

$$\eta_{mls,i} = \frac{\partial^2 L_i}{\partial w_i^2} w_i \bigg/ \frac{\partial L_i}{\partial w_i} \leq 0.$$

With concave revenue in q_i , quasiconcavity of profits clearly obtains while the labor supply function is concave, since this implies that revenue is also concave in w_i and that the cost function $w_i L_i$ is convex in w_i . Concavity of the labor supply function implies that the wage elasticity of marginal labor supply η_{mls} is negative and, therefore, condition (A.8) clearly holds. More generally, condition (A.8) holds if η_{mls} is not too large relative to η . Moreover, condition (A.8) holds irrespective of the curvature of the labor supply function if $\frac{L_i}{L_i - \alpha} = \frac{\alpha + \beta q}{\beta q}$, the inverse degree of economies of scale, is large. This is always the case for low levels of the marginal cost β . To again more intuition about the role of β , note that $1/\beta$ is the derivative of q_i with respect to L_i , hence for any given level, slope, and curvature of L_i , β determines the relative weight of the curvature of revenue on the curvature of the profit function. Hence, whenever $\frac{\partial L_i}{\partial w_i} \exists$, quasiconcavity obtains if β is sufficiently small or if the labor supply function is concave.

Whenever firm i chooses a wage so that its neighbor $i + j$ is just overbid and i starts competing with the next relevant competitor $i + j'$, $j' > j$, the labor supply function and thus the profit function exhibits a kink and $\frac{\partial L_i}{\partial w_i} \nexists$. However, we can show that the labor supply function is always flatter after the kink and hence, the kinks do not impair the concavity of the profit function. Let $\tilde{w}_{i,c} := \tilde{w}_{i,c}[\mathbf{w}_{-i}, \mathbf{m}_i]$ denote the wage where the relevant competitor on side $c = \ell, r$ is just overbid. Then, (A.4) implies that

$$\lim_{w_i \rightarrow \tilde{w}_{i,c}^-} \frac{\partial L_i^c}{\partial w_i} = \frac{L}{2H - \tilde{w}_{i,c} f'[d_{i,c}[\tilde{w}_{i,c}]] - w_{i+j} f'[2m_{i,i+j} - d_{i,c}[\tilde{w}_{i,c}]]} \frac{f[d_{i,c}[\tilde{w}_{i,c}]]^2}{\quad} \quad \text{and} \quad (\text{A.9})$$

$$\lim_{w_i \rightarrow \tilde{w}_{i,c}^+} \frac{\partial L_i^c}{\partial w_i} = \frac{L}{2H - \tilde{w}_i f'[d_{i,c}[\tilde{w}_{i,c}]] - w_{i+j'} f'[2m_{i,i+j'} - d_{i,c}[\tilde{w}_{i,c}]]} \frac{f[d_{i,c}[\tilde{w}_{i,c}]]^2}{\quad}. \quad (\text{A.10})$$

If $i + j$ was overbid by $i + j'$ at its own location, $\lim_{w_i \rightarrow \tilde{w}_{i,c}^-} \frac{\partial L_i^c}{\partial w_i} \geq \lim_{w_i \rightarrow \tilde{w}_{i,c}^+} \frac{\partial L_i^c}{\partial w_i}$ follows from $-f'[2m_{i,i+j'} - d_{i,c}[\tilde{w}_{i,c}]] > -f'[2m_{i,i+j} - d_{i,c}[\tilde{w}_{i,c}]] > 0$ and $w_{i+j'} > w_{i+j}$.

If $i + j$ was overbid by i at its own location, $\lim_{w_i \rightarrow \tilde{w}_{i,c}^-} \frac{\partial L_i^c}{\partial w_i} \geq \lim_{w_i \rightarrow \tilde{w}_{i,c}^+} \frac{\partial L_i^c}{\partial w_i}$ follows from $2m_{i,i+j'} - d_{i,c}[\tilde{w}_{i,c}] < 0$ and $-f'[2m_{i,i+j'} - d_{i,c}[\tilde{w}_{i,c}]] > 0 > -f'[2m_{i,i+j} - d_{i,c}[\tilde{w}_{i,c}]] > 0$.

A similar argument applies to the slope of the labor supply on firm i 's other side if it also exhibits a kink at $\tilde{w}_{i,c}$. Otherwise, the derivative of the labor supply function on that side exists. It follows that $\lim_{w_i \rightarrow \tilde{w}_{i,c}^-} \frac{\partial \pi_i}{\partial w_i} \geq \lim_{w_i \rightarrow \tilde{w}_{i,c}^+} \frac{\partial \pi_i}{\partial w_i}$ if $\left(\frac{mr_i}{\beta} - w_i\right) > 0$. If

$\left(\frac{mr_i}{\beta} - w_i\right) < 0$, then $\lim_{w_i \rightarrow \tilde{w}_{i,c}^-} \frac{\partial \pi_i}{\partial w_i}, \lim_{w_i \rightarrow \tilde{w}^+} \frac{\partial \pi_i}{\partial w_i} < 0$.

This proves that under the conditions specified above, profits are globally quasiconcave.

ii) The strategy space is convex and compact, and all solutions are interior if the degree of substitutability of products γ is sufficiently small. Firm i 's strategy space is given by the interval $S_i = [\underline{w}_{\alpha,i}, \bar{w}]$ and hence it is convex, closed and bounded. Interior solutions require that the slopes of the profit functions at the boundaries of the strategy space point inwards. At the lower bound, this condition holds if marginal revenue at \bar{w}_i , that is, at $q_i = 0$, is sufficiently large. Sufficiently small values of γ for any fixed number of firms, a given average price, fixed and variable cost, and labor market conditions $\mathbf{w}_{-i}, \mathbf{m}_i$, ensure that $\left. \frac{\partial \pi_i}{\partial w_i} \right|_{w_i=\underline{w}_i} = \left(\frac{p_{choke}}{\beta} - \underline{w}_i \right) \left. \frac{\partial L_i}{\partial w_i} \right|_{w_i=\underline{w}_i} - \alpha > 0$. $L_i[\bar{w}, \mathbf{w}_{-i}, \mathbf{m}_i] > \alpha$ if $\bar{w} > w_{\alpha,i}$ implies $\left. \frac{\partial \pi_i}{\partial w_i} \right|_{w_i=\bar{w}_i} < 0$.

iii) The Hessian of the log-transformed game is negative definite at the equilibrium point. We prove negative definiteness of the Hessian by showing that the game in transformed strategies $\ln w_i \in \tilde{S}_i$, where $\tilde{S}_i = [\ln w_{\alpha,i}, \ln \bar{w}]$, exhibits diagonal dominance at the equilibrium point where

$$\frac{\partial \pi_i}{\partial \ln w_i} = 0 \quad \forall i$$

Diagonal dominance at the equilibrium point requires that

$$\left| \frac{\partial^2 \pi_i}{\partial \ln w_i^2} \right| - \sum_{j \neq i}^N \left| \frac{\partial^2 \pi_i}{\partial \ln w_i \partial \ln w_j} \right| > 0, \quad (\text{A.11})$$

where

$$\begin{aligned} \frac{\partial^2 \pi_i}{\partial \ln w_i^2} &= w_i \frac{\partial \pi_i}{\partial w_i} + w_i^2 \frac{\partial^2 \pi_i}{\partial w_i^2} = w_i^2 \frac{\partial^2 \pi_i}{\partial w_i^2} \quad \text{since} \quad \frac{w_i \partial \pi_i}{\partial w_i} = 0 \\ \frac{\partial^2 \pi_i}{\partial \ln w_i \partial \ln w_j} &= \begin{cases} w_i w_j \frac{\partial^2 \pi_i}{\partial w_i \partial w_j} & \text{for } j = \ell, r \\ 0 & \text{for } j \neq i, \ell, r \end{cases} \end{aligned}$$

In an interior equilibrium no firm is overbid. This implies that around the equilibrium point the labor supply function is smooth and we do not need to worry about the kinks. Moreover, it implies that firm i 's relevant competitors are its immediate neighbors, that is, $\ell = i - 1$ and $r = i + 1$.

Firm i 's marginal profits change in its neighbors' log wages according to

$$\frac{\partial^2 \pi_i}{\partial w_i \partial w_c} w_i w_c = \frac{\partial m r_i}{\partial q_i} \frac{1}{\beta^2} \frac{\partial L_i}{\partial w_i} \frac{\partial L_i}{\partial w_c} w_i w_c - \frac{\partial L_i}{\partial w_c} w_i w_c + \left(\frac{m r_i}{\beta} - w_i \right) \frac{\partial^2 L_i}{\partial w_i \partial w_c} w_i w_c \leq 0 \quad (\text{A.12})$$

where

$$\frac{\partial L_i}{\partial w_c} w_c = \frac{L}{2H} f[d_{i,c}] \frac{dd_{i,c}}{dw_c} w_c < 0 \quad \text{since} \quad \frac{dd_{i,c}}{dw_c} = \frac{f[2m - d_{i,c}]}{w_i f'[d_{i,c}] + w_c f'[2m_{i,c} - d_{i,c}]} < 0 \quad (\text{A.13})$$

$$\begin{aligned} \frac{\partial^2 L_i}{\partial w_i \partial w_c} w_i w_c &= \frac{L}{2H} \left(2f'[d_{i,c}] \frac{dd_{i,c}}{dw_i} w_i \frac{dd_{i,c}}{dw_c} w_c + f'[2m_{i,c} - d_{i,c}] \frac{dd_{i,c}}{dw_i} w_i \frac{dd_{i,c}}{dw_i} w_c \right. \\ &\quad \left. + (w_i f''[d_{i,c}] - w_c f''[2m_{i,c} - d_{i,c}]) \frac{dd_{i,c}}{dw_i} \frac{dd_{i,c}}{dw_i} w_i \frac{dd_{i,c}}{dw_c} w_c \right) \\ &= \frac{L}{2H} \left(-2f'[d_{i,c}] \left(\frac{dd_{i,c}}{dw_i} w_i \right)^2 + f'[2m_{i,c} - d_{i,c}] \frac{dd_{i,c}}{dw_i} w_i \frac{dd_{i,c}}{dw_i} w_c \right. \\ &\quad \left. - (w_i f''[d_{i,c}] - w_c f''[2m_{i,c} - d_{i,c}]) \frac{dd_{i,c}}{dw_i} \left(\frac{dd_{i,c}}{dw_i} w_i \right)^2 \right) \leq 0. \quad (\text{A.14}) \end{aligned}$$

The second step follows from $\frac{dd_{i,c}}{dw_c} w_c = \frac{dd_{i,c}}{dw_i} w_i$. Since the first two terms in Equation (A.14) are always positive, it holds that

$$\left| \frac{\partial^2 \pi_i}{\partial w_i \partial w_c} w_i w_c \right| \leq \frac{\partial m r_i}{\partial q_i} \frac{1}{\beta^2} \frac{\partial L_i}{\partial w_i} \frac{\partial L_i}{\partial w_c} w_i w_c - \frac{\partial L_i}{\partial w_c} w_i w_c + \left| \left(\frac{m r_i}{\beta} - w_i \right) \frac{\partial^2 L_i}{\partial w_i \partial w_c} w_i w_c \right|.$$

Provided that condition i) holds, diagonal dominance as defined in Equation (A.11) at

the equilibrium point obtains if

$$\begin{aligned}
& -\frac{\partial m r_i}{\partial q_i} \frac{1}{\beta^2} \left(\frac{\partial L_i}{\partial w_i} w_i \right)^2 + 2 \frac{\partial L_i}{\partial w_i} w_i^2 - \left(\frac{m r_i}{\beta} - w_i \right) \frac{\partial^2 L_i}{\partial w_i^2} w_i^2 \\
& \geq \sum_{c=\ell, r} \left(\frac{\partial m r_i}{\partial q_i} \frac{1}{\beta^2} \frac{\partial L_i}{\partial w_i} \frac{\partial L_i}{\partial w_c} w_i w_c - \frac{\partial L_i}{\partial w_c} w_i w_c + \left(\frac{m r_i}{\beta} - w_i \right) \frac{\partial^2 L_i}{\partial w_i \partial w_c} w_i w_c \right) \\
\Leftrightarrow & -\frac{\partial m r_i}{\partial q_i} \frac{1}{\beta^2} \left(\frac{\partial L_i}{\partial w_i} w_i \right) \left(\frac{\partial L_i}{\partial w_i} w_i + \sum_{c=\ell, r} \frac{\partial L_i}{\partial w_c} w_c \right) + \frac{\partial L_i}{\partial w_i} w_i^2 + w_i \left(\frac{\partial L_i}{\partial w_i} w_i + \sum_{c=\ell, r} \frac{\partial L_i}{\partial w_c} w_c \right) \\
& - \left(\frac{m r_i}{\beta} - w_i \right) \left(\frac{\partial^2 L_i}{\partial w_i^2} w_i^2 - \sum_{c=\ell, r} \frac{\partial^2 L_i}{\partial w_i \partial w_c} w_i w_c \right) \geq 0 \\
\Leftrightarrow & \quad \frac{m r_i}{\beta} \frac{\partial L_i}{\partial w_i} w_i \geq 0
\end{aligned}$$

The last step follows from

$$\frac{\partial L_i}{\partial w_i} w_i = - \sum_{c=\ell, r} \frac{\partial L_i}{\partial w_c} w_c \quad \text{and} \quad \frac{\partial^2 L_i}{\partial w_i^2} w_i^2 - \sum_{c=\ell, r} \frac{\partial^2 L_i}{\partial w_i \partial w_c} w_i w_c = - \frac{\partial L_i}{\partial w_i} w_i. \quad (\text{A.15})$$

This proves that the game in log transformed strategies exhibits diagonal dominance at the equilibrium point.

Since conditions i) and ii) clearly also hold for the transformed game, there exists a unique solution to this game. Since $\ln \mathbf{w}$ is a monotone transformation of \mathbf{w} for $\mathbf{w} > 0$, uniqueness in the transformed game implies uniqueness of equilibrium in the original game. This completes the proof.

A.3 Proof of existence and uniqueness of the entry equilibrium (Lemma 2)

We first prove that under condition ii) of Lemma 2 there is a unique number of firms $N^* \geq 1$ corresponding to a (symmetric) second stage equilibrium that yields $\pi^*[\mathbf{m}_{sym}, N^*] = 0$. Moreover, we show that $\pi^*[\mathbf{m}_{sym}, N] = 0$ is decreasing in N .

Existence and uniqueness of the second-stage zero-profit equilibrium for symmetric distance patterns. As described in Section 2.3, for symmetric distance patterns a second-stage zero-profit equilibrium is given by a root of the function

$$G[m] := \pi^*[\mathbf{m}_{sym}, N] = \beta \rho[m] \psi[m] - \beta g[m] \quad (\text{A.16})$$

where $g[m] > 1$ is the inverse of an aggregate version of the familiar measure of the degree of scale economies, i.e., the ratio of average to marginal cost. We expect this to be falling in m : The larger firm size m , and the smaller the number of firms, the closer average cost to marginal cost. In turn, $\rho[m] := 1 + \frac{1}{\gamma N[m]}$ and $\psi[m] := \frac{\eta[m]+1}{\eta[m]}$ are the two markups on the goods and the labor market, respectively. Given that a symmetric equilibrium has $N = H/m$, we have $\rho_m = 1/(\gamma H) > 0$. As shown in Section 2.3, $\eta_m < 0$, whence we have $\psi_m = -\eta_m/\eta[m]^2 > 0$. As expected from intuition, both markups are falling in the number of firms and thus rising in the half-distance between two neighboring firms, m . Note that $G[m] > 0$ implies positive profits, while $G[m] < 0$ implies that firms make losses.

The following conditions are *sufficient* for a symmetric zero-profit equilibrium to *exist* and to be *unique*: a) $G[H] > 0$, b) $G[m]$ is continuous and $G_m > 0$ in the interval $(\tilde{m}, H]$, where \tilde{m} is defined by $\frac{L}{N[\tilde{m}]} \theta[\tilde{m}] = \alpha$.

Condition a) requires that a single firm in the market makes at least zero profits, that is,

$$\frac{L\theta[H]}{L\theta[H] - \alpha} \leq \left(1 + \frac{1}{\gamma}\right) \psi[H]. \quad (\text{A.17})$$

Observing that $\psi[m]$ increases in m , we can set $\psi[H]$ on the right-hand side to its minimum level of unity to obtain

$$\frac{\alpha}{L\beta}(1 + \gamma) \leq \frac{F[H]}{H} \quad (\text{A.18})$$

which is a sufficient condition for (A.17). It shows, that given α, β, L and H , the degree of substitutability of goods in the utility function γ , that governs the price elasticity of demand, must not be too large. Relating back to (A.17) in its original form, these restrictions imply that the price markup over marginal cost that a single firm can choose exceeds its average cost.⁴⁶

Condition b) requires that firm entry, which is associated with a decrease in the skill-reach m , lowers profits in the relevant range where firms produce positive output, that is, for $m \in (\tilde{m}, H]$. Since we know from above that $\rho_m > 0$ as well as $\psi_m > 0$, condition b) is satisfied if $g_m < 0$. It is straightforward to show that

$$g_m[m] = \frac{\frac{L}{H} f[m]}{\frac{mL}{H} \theta[m] - \alpha} \left(1 - \frac{\frac{mL}{H} \theta[m]}{\frac{mL}{H} \theta[m] - \alpha}\right) < 0 \quad \text{for } m \in (\tilde{m}, H]. \quad (\text{A.19})$$

⁴⁶ This condition is well known from the standard New Trade Theory model with homogeneous workers (cp. Equation (10) in Krugman, 1980).

Hence, there exists a unique $N^* \geq 1$ satisfying $G[m] = \pi^*[\mathbf{m}_{sym}, N] = 0$. Condition b) and $m = \frac{2H}{N}$ imply $\frac{\partial \pi^*[\mathbf{m}_{sym}, N]}{\partial N} < 0$.

Existence and uniqueness of the symmetric sub-game perfect equilibrium

With firm entry determined by (13) and consistent beliefs as described in Section (2.2), existence of N^* satisfying $\pi^*[\mathbf{m}_{sym}, N^*] = 0$ implies that no (symmetric or asymmetric) distance pattern involving a number of firms smaller N^* can be an equilibrium according to (15).

To prove that $\mathbf{m}_{sym}|N^*$ is the unique equilibrium, it remains to show that there is no other distance pattern with $N \geq N^*$ that is consistent with (13) and (15). We do so by showing that every *asymmetric* distance pattern with $N \geq N^*$ firms and every *symmetric* distance pattern with $N > N^*$ firms implies negative profits for at least one firm, which implies $\nu(N) = 0$ for all $N > N^*$. The result for symmetric distance vectors follows readily from $\frac{\partial \pi^*[\mathbf{m}_{sym}, N]}{\partial N} < 0$. The proof for asymmetric location patterns is slightly more involved and requires restrictions on the parameter space. It runs along the following line of argument. We conjecture that the symmetric zero profit solution characterized by $\pi^*[\mathbf{m}_{sym}, N^*] = 0$ is an equilibrium and then consider any possible change towards an asymmetric location pattern featuring the same or a larger number of firms. Since the symmetric location pattern maximizes labor supply per firm, such a change must bring about a decrease in q for at least one firm. Let j be the firm that produces the smallest quantity in any arbitrarily chosen asymmetric location pattern with N^* firms. Then, we can show that if the marginal cost β are sufficiently small, firm j 's profits must be negative in the asymmetric pattern. Hence, no asymmetric location pattern can be an equilibrium consistent with optimal entry choices of entrants and non-entrants.

From the point of view of any firm i , a zero-profit equilibrium is characterized by

$$\rho[N, q_i, \mathbf{q}_{-i}] \cdot \psi[\mathbf{m}_i, N] = g[q_i] \quad \text{where} \quad g[q_i] = \frac{L_i}{L_i - \alpha} = \frac{\alpha + \beta q_i}{\beta q_i}, \quad \forall$$

which states that the product of markups equals the inverse of the degree of economies of scale (cp. Equation (23)).⁴⁷ Note that $q_i = \frac{L_i[\mathbf{m}_i, N] - \alpha}{\beta}$ is also a function of the location pattern and so are $\rho[N, q_i, \mathbf{q}_{-i}]$ and $g[q_i]$.⁴⁸ However, it will prove important that $g[\cdot]$ and $\rho[\cdot]$ depend on \mathbf{m}_i *only* through output quantities, since we may therefore pin down the

⁴⁷ Using the first-order condition (11), we can write optimum profits $\pi^*[\mathbf{m}_i, N]$ of any firm as

$$\pi_i^* = w_i^* L_i \left(\psi_i \rho_i \frac{L_i - \alpha}{L_i} - 1 \right).$$

Then, for $w_i^*, L_i > 0$, $\pi_i^* = 0$ iff $\rho_i \psi_i = g_i$.

⁴⁸ The dependence of ρ_i on \mathbf{q}_{-i} derives from the dependence of ρ_i on total expenditure $Y = \sum_k^N r_i[q_k]$.

changes in $g[\cdot]$ and $\rho[\cdot]$ for firm j for an arbitrary change in the location pattern, since, by definition, q_j decreases whenever we move away from symmetry.

From the point of view of any firm, the change towards an asymmetric location pattern can be described in terms of changes in potentially all elements of the distance vector \mathbf{m}_i . Its profits are affected by corresponding changes in the markups and the degree of economies of scale. In the new location pattern, firm i 's profits will be negative if and only if the total markup ($\rho_i \psi_i$) increases by less (falls by more) than the inverse of the degree of economies of scale, g_i . That is, a sufficient condition for $\pi^*[\mathbf{m}_i, N^*] < 0$ for all $\mathbf{m}_i \neq \mathbf{m}_{sym}$ is that

$$d(\rho_i \cdot \psi_i) = \rho_i d\psi_i + \psi_i d\rho_i < dg_i, \quad \text{where} \quad (\text{A.20})$$

$$dg_i = \frac{\partial g[q_i]}{\partial q_i} \frac{\partial q_i}{\partial L_i} dL_i = -\frac{\alpha}{(\beta q_i)^2} dL_i \quad \text{with} \quad dL_i = \sum_k^N \frac{\partial L_i}{\partial m_{k,k+1}} dm_{k,k+1} \quad (\text{A.21})$$

$$d\rho_i = \sum_k^N \frac{\partial \rho_i}{\partial q_k} \frac{\partial q_k}{\partial L_k} dL_k \quad (\text{A.22})$$

$$d\psi_i = \frac{1}{\eta_i} \frac{dL_i}{L_i} - \frac{1}{\eta_i} \sum_k^N \frac{\frac{\partial^2 L_i}{\partial w_i \partial m_{k,k+1}}}{\frac{\partial L_i}{\partial w_i}} dm_{k,k+1} \quad (\text{A.23})$$

Note that L_i depends on the distances between *all* firms and not just firm i 's neighbors, because it is a function of firm i 's own and its neighbors' equilibrium wages in the second stage, which jointly solve all firms' first-order conditions and therefore depend on the complete distance vector.

Consider the problem of firm j , defined as the firm that produces the smallest amount of output in the asymmetric allocation. Then it is true that $dL_j < 0$ and $dg_j > 0$. Moreover, in the asymmetric allocation it is true that $q_j \leq \frac{1}{N} \sum_k q_k$, which, by concavity of the revenue function, implies $\ln p_j \geq \overline{\ln p}$. Therefore, it follows from $\rho_j = 1 + \frac{\delta_j}{\gamma}$ that

$$d\rho_j = \frac{1}{\gamma} (\delta_j^{asym} - \delta_j^{sym}) = \overline{\ln p} - \ln p_i \leq 0.$$

Firm j 's price markup weakly decreases because its expenditure share is weakly smaller in the asymmetric location pattern. Note that an asymmetric location pattern where all firms produce the same quantity is conceivable. In this situations, the above statements hold with equality. The decrease in the degree of economies of scale $dg_j > 0$ and the decrease in the price markup work towards lowering firm j 's optimum profits. However, the effect of the change in the location pattern on the wage markup is ambiguous. The first term is strictly negative for $dL_j < 0$, but the second term, reflecting the sum of the elasticities of marginal labor supply with respect to the change in the location pattern, is difficult to

sign. It represents the change in competitiveness of firm j 's labor market environment due to changes in the distances to its neighbors and the equilibrium wage adjustments to the change in the overall distance pattern. Therefore, according to Equation (A.21), it holds that firm j 's optimum profits decrease whenever the sum of the effects on the degree of economies of scale, the price markup, and negative effect on the wage markup due $dL_j < 0$ overcompensate a potentially positive effect on the wage markup due to a decrease in the degree of competitiveness of firm j 's labor market environment. This is always true if the marginal cost β are small relative to the fixed cost α , as then, the effect on average cost is large compared to the adjustment in the wage markup, which is independent of β . Hence, for sufficiently small β it holds that every departure from symmetry (holding fixed N) leads to a decrease in firm j 's profits. Since we are starting from the zero-profit equilibrium, firm j 's profits will be negative in any asymmetric location pattern featuring the same number of firms as the symmetric starting point.

Thus, no asymmetric location pattern with $N = N^*$ exists where all firms make positive profits. Moreover, the exact same rationale implies that no asymmetric pattern with a number of firms larger N^* exists where all firms make positive profits. This completes the proof.

A.4 The limiting case of $H \rightarrow 0$

As we let the degree of skill heterogeneity approach zero, our equilibrium converges to the equilibrium of a monopolistic competition model with translog preferences. From the previous appendix it follows that if an equilibrium exists with some \bar{H} , it also exists for $H < \bar{H}$. In all of these equilibria, m will be smaller than \bar{H} , ensuring $H/m = N > 1$. Consider an exogenous decrease in the degree of skill differentiation $\hat{H} < 0$ within the interval $(0, \bar{H}]$. A smaller circumference means that the mass of labor on any interval of the skill circle increases. Holding m constant for a moment, this would allow firms to expand output without having to rely on workers with less suitable types of skills, thus increasing the degree of scale economies and decreasing $g[m]$. Moreover, from $N = H/m$ a smaller H means a lower number of firms, which implies a higher goods price markup. But this, together with the size effect, implies positive profits. Hence, $\hat{N} = \hat{H}$ with $\hat{m} = 0$ is not an equilibrium adjustment. Totally differentiating (23), we obtain

$$\hat{m} = \frac{g_H - \psi[m]\rho_H}{-g_m + \psi[m]\rho_m + \rho[m]\psi_m} \frac{H}{m} \hat{H} = \frac{g[m](g[m] - 1) + \psi[m]\frac{m}{\gamma H}}{\frac{f[m]}{\theta[m]}g[m](g[m] - 1) + \psi[m]\frac{m}{\gamma H} + \frac{\psi_m m}{\psi[m]}} \hat{H}. \quad (\text{A.24})$$

The “multiplier” in front of \hat{H} is positive, meaning that m falls as H decreases, but $f[m]/\theta[m] < 1$ and $\psi_m m/\psi[m] \geq 0$ imply that the multiplier can be greater or smaller one. Thus, the net effect on $N = H/m$ is generally ambiguous. Now, let $H \rightarrow 0$,

whence $m = H/N$ must approach zero as well. Therefore, $f[m]/\theta[m]$ goes to unity and $\psi_m m/\psi[m] \geq 0$ goes to zero, so that the multiplier approaches unity and N converges to a constant \underline{N} . Returning to the equilibrium condition (23) and letting $m \rightarrow 0$ ($\theta[m] \rightarrow 1$, $\psi[m] \rightarrow 1$) and $H/m = N \rightarrow \underline{N}$. We finally obtain that \underline{N} must satisfy

$$\frac{L}{L - \alpha \underline{N}} = 1 + \frac{1}{\gamma \underline{N}} \quad (\text{A.25})$$

which is the equilibrium condition for the number of firms in a Krugman (1979)-type model with homogeneous workers and translog preferences.

A.5 The constrained social optimum

The social planner maximizes log utility with respect to m and subject to the condition that price equals average cost (AC) and the endowment constraint which we can combine to $p = \frac{L\theta[m]}{L\theta[m] - \alpha N[m]}$:

$$\max_m \ln V = \ln \theta[m] - \left(\frac{1}{2\gamma N[m]} + \ln p[m] \right) \quad \text{s.t.} \quad p[m] = \frac{L\theta[m]}{L\theta[m] - \alpha N[m]} \quad (\text{A.26})$$

The first order condition results as

$$\frac{Lf[m]}{L\theta[m] - \frac{\alpha H}{m}} = 1 + \frac{m}{2\gamma H}. \quad (\text{A.27})$$

The second order condition for a maximum holds since, as we can show, the welfare function is globally concave, i.e.

$$\frac{d^2 \ln V}{dm^2} = -\frac{\left(L\theta_m[m] + \frac{\alpha H}{m^2}\right)^2}{\left(L\theta[m] - \frac{\alpha H}{m}\right)^2} + \frac{L\theta_{mm}[m] - 2\frac{\alpha H}{m^3}}{L\theta[m] - \frac{\alpha H}{m}} < 0. \quad (\text{A.28})$$

A sufficient condition for this to hold is

$$\theta_{mm}[m] := \frac{\partial^2 \theta[m]}{\partial m^2} = \frac{1}{m} \left(f'[m] - \frac{2}{m} f[m] + \frac{2}{m} \theta[m] \right) \leq 0 \quad (\text{A.29})$$

which requires $f[m] \geq \theta[m] + \frac{m}{2} f'[m]$. Since concavity of $f[\cdot]$ implies $f[m] \geq f\left[\frac{m}{2}\right] + \frac{m}{2} f'[m]$ and (by Jensen's inequality) $f\left[\frac{m}{2}\right] \geq \theta[m]$, it follows that $f[m] \geq f\left[\frac{m}{2}\right] + \frac{m}{2} f'[m] \geq \theta[m] + \frac{m}{2} f'[m]$ and therefore $\theta_{mm}[m] \leq 0$ and $\frac{\partial^2 \ln V}{\partial m^2} < 0$ always hold.

To compare the planner's solution with the laissez faire equilibrium determined by (23)

we rewrite (A.27) as

$$g[m] = \frac{\theta[m]}{f[m]} \frac{1}{\psi[m]} \psi[m] \rho[m/2]. \quad (\text{A.30})$$

The difference between the two conditions appears on the right-hand side of this equation. Since $g_m < 0$, the social planner's solution implies a larger m than the market equilibrium, if the right-hand side is smaller than $\psi[m]\rho[m]$ for all values of m . Since $\rho_m > 0$,

$$\frac{\theta[m]}{f[m]} \frac{1}{\psi[m]} < 1 \quad (\text{A.31})$$

is a sufficient condition for this to hold. We show next that concavity of $f[\cdot]$ suffices to establish this result. Rearranging (A.31) and inserting $\psi[m] = \frac{f[m]^2 - 2f'[m]m\theta[m]}{f[m]^2}$ yields $\frac{1 + \frac{2}{f[m]}f'[m]m}{f[m]} < \frac{1}{\theta}$ which holds a fortiori because concavity of $f[\cdot]$ implies that $\frac{1 + f'[m]m}{f[m]} < 1$. Hence, condition (A.31) is fulfilled and it follows that the market equilibrium firm size is too small compared to the socially optimal allocation.

A.6 Further details of the trading equilibrium

A.6.1 Proof of Proposition 1

(i) Log-differentiating the equilibrium condition (27) and setting $k = 1$, we obtain

$$\hat{m} = A \cdot \hat{k} \quad \text{with} \quad A := \frac{\psi[m] \frac{1}{\gamma H}}{-g_m[m] + \psi[m] \frac{1}{\gamma H} + \rho^T[m] \psi_m[m]}. \quad (\text{A.32})$$

Since $g_m < 0$ and $\psi_m > 0$,⁴⁹ we find that $0 < A < 1$ which implies $0 < \hat{m} = A \cdot \hat{k} < \hat{k}$. Hence, m increases and the number of firms in each country falls. However, $A < 1$ implies that the total number of available varieties $N^T = k \cdot N > N^A$ is still larger with trade than under autarky.

(ii) As the price markup depends negatively on the number of available varieties $k \cdot N$, it follows directly from the previous result that it must fall. Furthermore, we know from above that the wage markup increases. Log-differentiating (25) and again setting $k = 1$

⁴⁹see Appendix A.3 for details

yields

$$\hat{p} = B \cdot \hat{k} \quad \text{with} \quad B = \frac{\frac{m}{\gamma H}}{\left(1 + \frac{m}{\gamma H}\right) \left(-g_m[m] + \psi[m] \frac{1}{\gamma H} + \rho^T[m] \psi_m[m]\right)}. \quad (\text{A.33})$$

Since $-1 < B < 0$, it follows that $\hat{p} < 0$.

(iii) This follows from $\theta_m = \frac{1}{m} (f[m] - \theta[m]) < 0$.

(vi) Real income, $\theta[m]/p[m]$, must increase by virtue of the excess entry result demonstrated in A.5. With higher real income and a larger variety available for consumption as established in (i), it follows from (24) that welfare of the worker earning average income increases.

A.6.2 The first order conditions with two symmetric countries and positive trade cost

Under the assumption that the constraints $q_i, q_i^* \geq 0$ never bind, we may write (28) as

$$\max_{w_i, q_i} \left\{ r_i[q_i, N, \overline{\ln p}, Y] + r_i^* \left[\frac{\bar{q}_i - q_i}{\tau}, N, \overline{\ln p}, Y \right] - w_i L_i \right\}.$$

The first order condition with respect to w_i then obtains as

$$\frac{p^*}{\tau} \left(\frac{\partial \ln p^*}{\partial \ln \frac{\bar{q}_i - q_i}{\tau}} + 1 \right) \frac{\partial L_i}{\partial w_i} = w_i \frac{\partial L_i}{\partial w_i} + L_i \quad \Leftrightarrow \quad p^* = \frac{\varepsilon_i^*}{\varepsilon_i^* - 1} \frac{\eta_i + 1}{\eta_i} w_i \tau$$

and the first order condition with respect to q_i reads

$$p \left(\frac{\partial \ln p}{\partial \ln q_i} + 1 \right) \frac{\partial L_i}{\partial w_i} = \frac{p^*}{\tau} \left(\frac{\partial \ln p^*}{\partial \ln \frac{\bar{q}_i - q_i}{\tau}} + 1 \right) \frac{\partial L_i}{\partial w_i} \quad \Leftrightarrow \quad p \frac{\varepsilon_i - 1}{\varepsilon_i} = \frac{p^*}{\tau} \frac{\varepsilon_i^* - 1}{\varepsilon_i^*}.$$

Both first order conditions together imply (31) and (32).

A.6.3 Proof of Proposition 2

In the symmetric equilibrium with identical countries the average price in the domestic and the foreign market is the same and given by $\overline{\ln p} = \overline{\ln p^*} = 1/2 \ln p + 1/2 \ln p^*$. Inserting $\overline{\ln p}$ and $\overline{\ln p^*}$ into the Z-terms in (31), (32), we can use the same logic as in A.1 to obtain explicit solutions for p and p^* , where the price markups no longer depend on the own

price, but only on the respective other price and the number of firms:

$$p = \frac{\mathcal{W}[\tilde{Z}]}{2}\psi \quad \text{with } \tilde{Z} = \frac{2}{\psi} \exp \left\{ 2 + \frac{m}{\gamma H} + \ln p^* \right\} \quad (\text{A.34})$$

$$p^* = \frac{\mathcal{W}[\tilde{Z}^*]}{2}\psi\tau \quad \text{with } \tilde{Z}^* = \frac{2}{\psi\tau} \exp \left\{ 2 + \frac{m}{\gamma H} + \ln p \right\}. \quad (\text{A.35})$$

Inserting $p = \frac{\mathcal{W}[\tilde{Z}]}{2}\psi$ and $p^* = \frac{\mathcal{W}[\tilde{Z}^*]}{2}\psi\tau$ into the \tilde{Z} -terms, we obtain

$$p = \mathcal{W} \left[\mathcal{W}[\tilde{Z}^*]\tau \exp \left\{ 2 + \frac{m}{\gamma H} \right\} \right] \frac{\psi}{2} \quad (\text{A.36})$$

$$p^* = \mathcal{W} \left[\frac{\mathcal{W}[\tilde{Z}]}{\tau} \exp \left\{ 2 + \frac{m}{\gamma H} \right\} \right] \frac{\psi}{2}\tau. \quad (\text{A.37})$$

It proves convenient to focus on the price markup values $W = \mathcal{W}[\tilde{Z}]$ and $W^* = \mathcal{W}[\tilde{Z}^*]$ instead of prices. The corresponding system of equations determining these values emerges as

$$W = W[W^*, m] = \mathcal{W} \left[W^*\tau \exp \left\{ 2 + \frac{m}{\gamma H} \right\} \right] \quad (\text{A.38})$$

$$W^* = W^*[W, m] = \mathcal{W} \left[\frac{W}{\tau} \exp \left\{ 2 + \frac{m}{\gamma H} \right\} \right]. \quad (\text{A.39})$$

Note that for zero trade costs ($\tau = 1$) the price markups are identical. While the markup on domestic varieties increases in τ , the markup on foreign varieties falls in the level of trade costs. For any $\tau > 1$, it must therefore be true that $W > W^*$.

Note that the two country version of (A.1) can be written as $p = \left(1 + \frac{1}{\gamma N^T} + \frac{1}{2} \ln p^* - \frac{1}{2} \ln p \right) \tilde{w}$ and analogously for p^* . In view of (A.34) and (A.35) it follows that $\frac{W}{2} = 1 + \frac{1}{\gamma N^T} + \frac{1}{2} \ln p^* - \frac{1}{2} \ln p$ and $\frac{W^*}{2} = 1 + \frac{1}{\gamma N^T} + \frac{1}{2} \ln p - \frac{1}{2} \ln p^*$. The expenditure shares in (29) can therefore be written as

$$\delta = \left(\frac{W}{2} - 1 \right) \gamma \quad \text{and} \quad \delta^* = \left(\frac{W^*}{2} - 1 \right) \gamma. \quad (\text{A.40})$$

Direct demand functions for foreign varieties in terms of W^* obtain as

$q^* = \frac{\delta^* Y}{p^*} \left(1 - \frac{2}{W^*} \right) \frac{\gamma Y}{\psi}$. This implies that the prohibitive level of trade costs $\bar{\tau}$ for which $q^* = 0$ satisfies $\mathcal{W} \left[\frac{W}{\bar{\tau}} \exp \left\{ 2 + \frac{2}{\gamma N^T} \right\} \right] \equiv 2$. It follows that for non-prohibitive trade costs $W \geq W^* \geq 2$. Inserting demand and income $Y = L\theta$ into the labor market clearing

condition (33), and rearranging terms gives

$$\begin{aligned}\gamma \left(2 - \frac{2}{W} - \frac{2}{W^*}\right) &= \frac{\frac{L\theta[m]}{N[m]} - \alpha}{L\theta[m]} \psi[m] \\ \gamma h[W, W^*] &= \frac{\psi[m]}{g[m]N[m]}.\end{aligned}\tag{A.41}$$

For easier reference the second line introduces $h[W, W^*] := (2 - \frac{2}{W} - \frac{2}{W^*})$. (A.41), (A.38) and (A.39) form our system of equations in W, W^* and m .

(i) Comparative statics of firm size and markups. The proof of proposition 2 requires that we solve this system for an exogenous change in τ . Doing so by log-linearization, we write the solution as $\widehat{W} = \omega \cdot \hat{\tau}$, $\widehat{W}^* = \omega^* \cdot \hat{\tau}$ and $\widehat{m} = \mu \cdot \hat{\tau}$. We next explore the sign of the elasticities ω, ω^* and μ . For notational convenience we suppress the functional dependence of N and ψ on m in the following, whenever it is not crucial. Log- differentiating (A.41), (A.38), (A.39) leads to

$$\begin{aligned}\begin{bmatrix} -\frac{\partial \ln h}{\partial \ln W} & -\frac{\partial \ln h}{\partial \ln W^*} & \frac{\partial \ln \psi}{\partial \ln m} - \frac{\partial \ln g}{\partial \ln W} - \frac{\partial \ln N}{\partial \ln m} \\ -1 & \frac{\partial \ln W^*}{\partial \ln W} & -1 \\ \frac{\partial \ln W^*}{\partial \ln W} & -1 & \frac{\partial \ln m}{\partial \ln W^*} \end{bmatrix} \begin{bmatrix} \widehat{W} \\ \widehat{W}^* \\ \widehat{m} \end{bmatrix} &= \begin{bmatrix} 0 \\ -\frac{\partial \ln W}{\partial \ln \tau} \cdot \hat{\tau} \\ -\frac{\partial \ln W^*}{\partial \ln \tau} \cdot \hat{\tau} \end{bmatrix} \\ \begin{bmatrix} -\frac{1}{W-1-\frac{W}{W^*}} & -\frac{1}{W^*-1-\frac{W^*}{W}} & \frac{\frac{Lf[m]}{N}}{\frac{L\theta[m]}{N}-\alpha} - \frac{f[m]}{\theta[m]} + 1 + \frac{\partial \ln \psi}{\partial \ln m} \\ -1 & \frac{1}{W+1} & \frac{1}{\gamma^N} \frac{1}{W+1} \\ \frac{1}{W^*+1} & -1 & \frac{1}{\gamma^N} \frac{1}{W^*+1} \end{bmatrix} \begin{bmatrix} \widehat{W} \\ \widehat{W}^* \\ \widehat{m} \end{bmatrix} &= \begin{bmatrix} 0 \\ -\frac{1}{W+1} \cdot \hat{\tau} \\ \frac{1}{W^*+1} \cdot \hat{\tau} \end{bmatrix}.\end{aligned}\tag{A.42}$$

Denoting the 3×3 -matrix of derivatives by D , it follows that

$$\omega = \frac{1}{(W+1)(W^*+1)} \left[\left(\frac{\frac{Lf[m]}{N}}{\frac{L\theta[m]}{N}-\alpha} - \frac{f[m]}{\theta[m]} + 1 + \frac{\partial \ln \psi}{\partial \ln m} \right) W^* - \frac{1}{\gamma N h[W, W^*]} \frac{4}{W^*} \right] \frac{1}{\det[D]}\tag{A.43}$$

$$\omega^* = \frac{1}{(W+1)(W^*+1)} \left[- \left(\frac{\frac{Lf[m]}{N}}{\frac{L\theta[m]}{N}-\alpha} - \frac{f[m]}{\theta[m]} + 1 + \frac{\partial \ln \psi}{\partial \ln m} \right) W + \frac{1}{\gamma N h[W, W^*]} \frac{4}{W} \right] \frac{1}{\det[D]}\tag{A.44}$$

$$\mu = \frac{2W^*/W - 2W/W^*}{h[W, W^*](W+1)(W^*+1)} \frac{1}{\det[D]}.\tag{A.45}$$

The signs of the elasticities hinge upon the sign of the determinant which is given by

$$\det[D] = \left(\frac{\frac{Lf[m]}{N}}{\frac{L\theta[m]}{N} - \alpha} - \frac{f[m]}{\theta[m]} + 1 + \frac{\partial \ln \psi}{\partial \ln m} \right) \frac{WW^* + W + W^*}{(W+1)(W^*+1)} - \frac{1}{\gamma Nh[W, W^*]} \frac{(2+W^*)\frac{2}{W} + (2+W)\frac{2}{W^*}}{(W+1)(W^*+1)}. \quad (\text{A.46})$$

Since $WW^* > 2$ and $W \geq W^*$, we have $WW^* + W + W^* > (2+W^*)\frac{2}{W} + (2+W)\frac{2}{W^*}$. This implies that $\det[D] > 0$ if

$$\frac{\frac{Lf[m]}{N}}{\frac{L\theta[m]}{N} - \alpha} - \frac{f[m]}{\theta[m]} + 1 + \frac{\partial \ln \psi}{\partial \ln m} > \frac{1}{\gamma Nh[W, W^*]}. \quad (\text{A.47})$$

We know from above that $\frac{f[m]}{\theta[m]} < 1$ and $\frac{\partial \ln \psi}{\partial \ln m} > 0$, and therefore, inequality (A.47) holds if

$$\frac{\frac{Lf[m]}{N}}{\frac{L\theta[m]}{N} - \alpha} > \frac{1}{\gamma Nh[W, W^*]}. \quad (\text{A.48})$$

Using the equilibrium condition (A.41), we can rewrite this as $\psi[m] \geq \theta[m]/f[m]$. We have proven in Appendix A.5 that this inequality always holds. Hence, it follows that $\det[D] > 0$.

Returning to our elasticity ω , we note that $W^* \geq \frac{4}{W^*}$, $\det[D] > 0$ and (A.47) imply $\omega > 0$. By analogy, it follows that $\omega^* < 0$. And finally, $W \geq W^*$ implies that $\mu \leq 0$. For reasons pointed out in the text, μ is monotonic in the initial level of trade costs, converging to zero as τ approaches one. Looking at A.45, the level of τ enters through W and W^* . The lower the trade cost level, the smaller the difference between W and W^* . At $\tau = 1$, price markups are identical and $m = 0$. This proves part (i) of the proposition.

(ii) Changes in prices. The proposition states that for $\hat{\tau} < 0$, $\hat{p}^* < 0$ while \hat{p} is ambiguous. The price of imported varieties is affected by the change in τ and the changes in both markups

$$\hat{p}^* = \left(\omega^* + \frac{\partial \ln \psi}{\partial \ln m} \mu + 1 \right) \hat{\tau} \quad (\text{A.49})$$

where $\frac{\partial \ln \psi}{\partial \ln m} = \frac{-2mf''[m]F[m]}{f[m]^2\psi[m]} - \frac{2mf'[m]}{f[m]} > 0$. Inserting (A.44) and (A.45) shows that \hat{p}^* is positive if and only if

$$-\frac{d_{13}W - \frac{2}{W}\frac{2}{\gamma h N} + \frac{\partial \ln \psi}{\partial \ln m} \frac{1}{h[W, W^*]} \left(\frac{2W}{W^*} - \frac{2W^*}{W} \right)}{d_{13}(WW^* + W + W^*) - \frac{2}{\gamma h[W, W^*]N} \left(\frac{2+W^*}{W} + \frac{2+W}{W^*} \right)} + 1 > 0 \quad (\text{A.50})$$

where d_{13} is the element in row 1 and column 3 of D . Canceling identical terms in the denominator and the numerator shows that this is true if $\frac{\frac{\partial \ln \psi}{\partial \ln m} \frac{1}{h[W, W^*]} \left(\frac{2W}{W^*} - \frac{2W^*}{W} \right)}{d_{13}(WW^* + W^*) - \frac{2}{\gamma h[W, W^*]N} \left(\frac{W^*}{W} + \frac{2+W}{W^*} \right)} <$

1. Noting that $d_{13} = \frac{\frac{Lf[m]}{N}}{\frac{L\theta[m]}{N} - \alpha} - \frac{f[m]}{\theta[m]} + 1 + \frac{\partial \ln \psi}{\partial \ln m}$ and observing the inequality in (A.48), it follows that $WW^* + W^* \geq \frac{2W^*}{W} + \frac{4+2W}{W^*}$ and $WW^* + W^* \geq \frac{1}{h[W, W^*]} \left(\frac{2W}{W^*} - \frac{2W^*}{W} \right)$ is sufficient for the inequality in (A.50) to hold. Using from above $W \geq W^* \geq 2$, it is straightforward to show that these two conditions are fulfilled.

The change in the domestic price obtains as

$$\hat{p} = \left(\omega + \frac{\partial \ln \psi}{\partial \ln m} \mu \right) \hat{\tau}. \quad (\text{A.51})$$

We know from above that $\omega > 0$; the pro-competitive effect of lower trade costs on the goods market. This is potentially offset by an increase in the wage markup. For τ close to one, the goods market effect clearly dominates as μ is close to zero.

Conversely, at $\bar{\tau}$ (prohibitive trade cost level), the labor market effect dominates. Inserting (A.43) and (A.45) gives

$$\hat{p} = \left[W^* \left(\frac{\frac{Lf[m]}{N}}{\frac{L\theta[m]}{N} - \alpha} - \frac{f[m]}{\theta[m]} + 1 + \frac{\partial \ln \psi}{\partial \ln m} \right) - \frac{2}{\gamma N h[W, W^*]} \frac{2}{W^*} - \frac{\partial \ln \psi}{\partial \ln m} \frac{1}{h[W, W^*]} \left(\frac{2W}{W^*} - \frac{2W^*}{W} \right) \right] \times \frac{1}{(W+1)(W^*+1) \det[D]} \hat{\tau}. \quad (\text{A.52})$$

Remember that prohibitive trade costs imply an infinite price elasticity and therefore a price markup of zero, whence $W^* = 2$. To see if $\hat{p} > 0$ for $\tau = \bar{\tau}$, as stated in proposition 2, we must therefore evaluate the bracketed term at $W^* = 2$. We obtain

$$-2 \frac{\frac{Lf[m]}{N}}{\frac{L\theta[m]}{N} - \alpha} + 2 \frac{f[m]}{\theta[m]} - 2 - 2 \frac{\partial \ln \psi}{\partial \ln m} + \frac{2}{\gamma N h[W, W^*]} \frac{2}{W^*} + \frac{\partial \ln \psi}{\partial \ln m} (W+2) \quad (\text{A.53})$$

Inserting the equilibrium condition (A.41), which reduces to $\gamma h[W, W^*] = \frac{L\theta[m]/N - \alpha}{L\theta[m]} \psi =$

$\frac{2}{W} \frac{1}{N}$ at $\tau = \bar{\tau}$, shows that the expression is negative, if

$$\psi W \frac{f[m]}{\theta[m]} < 2 \frac{f[m]}{\theta[m]} + W - 2 + W \frac{\partial \ln \psi}{\partial \ln m}. \quad (\text{A.54})$$

Inserting the explicit expressions for ψ and $\frac{d \ln \psi}{d \ln m}$ leads to

$$\frac{W}{\theta[m]} \frac{f[m]^2 - 2f'[m]F[m]}{f[m]} < W - 2 + \frac{2f[m]}{\theta[m]} + W \left(\frac{-2f''[m]\theta}{f[m]^2\psi} - \frac{2mf'[m]}{f[m]} \right). \quad (\text{A.55})$$

Since $f''[m] \leq 0$, the inequality holds if

$$\frac{W}{\theta[m]} \frac{f[m]^2 - 2f'[m]F[m]}{f[m]} < W - 2 + \frac{2f[m]}{\theta[m]} - W \frac{2mf'[m]}{f[m]}. \quad (\text{A.56})$$

Rearranging terms shows that this inequality holds if $f[m] < \theta[m]$, which is true given $f'[m] < 0$. This completes the proof of part (ii) of proposition 2.

(iii) Welfare. Indirect utility of the worker receiving average income in the equilibrium with trade costs is given by $\ln V = \ln \theta[m] - \ln P^T[p, p^*, m]$, where

$$\ln P^T[p, p^*, m] = \frac{1}{2\gamma N^T} + \frac{1}{N^T} \sum_{i=1}^{N^T} \ln p_i + \frac{\gamma}{2N^T} \sum_{i=1}^{N^T} \sum_{j=1}^{N^T} \ln p_i (\ln p_j - \ln p_i) \quad (\text{A.57})$$

with $N^T = N + N^*$ and i, j indexing domestic *and* foreign varieties. Under symmetry, which implies $N^* = N = N^T/2$, the price index simplifies to

$$\ln P^T[p, p^*, m] = \frac{1}{4\gamma N} + \frac{1}{2} \ln p + \frac{1}{2} \ln p^* - \frac{\gamma N}{4} (\ln p - \ln p^*)^2. \quad (\text{A.58})$$

The change in indirect utility is then

$$\widehat{V} = \left(\frac{\partial \ln \theta}{\partial \ln m} - \frac{\partial \ln P}{\partial \ln m} \right) \widehat{m} - \frac{\partial \ln P}{\partial \ln p} \widehat{p} - \frac{\partial \ln P}{\partial \ln p^*} \widehat{p}^* \quad (\text{A.59})$$

with $\frac{\partial \ln \theta}{\partial \ln m} = \frac{f[m] - \theta[m]}{\theta[m]} < 0$, $\frac{\partial \ln P}{\partial \ln m} = \frac{1}{4\gamma N} + \frac{\gamma N}{4} (\ln p - \ln p^*)^2 > 0$, $\frac{\partial \ln P}{\partial \ln p} = \frac{1}{2} + \frac{\gamma N}{2} (\ln p - \ln p^*) \geq 0$ and $\frac{\partial \ln P}{\partial \ln p^*} = N\delta^* \geq 0$. Inserting yields equation (34).

Using the results that at the prohibitive level of trade costs $\delta^* = 0$, $\widehat{p} > 0$ and $\widehat{m} > 0$, it follows from (34) that $\widehat{V} < 0$ at $\tau = \bar{\tau}$. Since at $\tau = 1$ it holds that $\widehat{m} = 0$, $\widehat{p} < 0$ and $\widehat{p}^* < 0$, it follows that $\widehat{V} > 0$ at $\tau = 1$.

A.7 Additional details of the trade and migration equilibrium

A.7.1 Conditions for existence and uniqueness of the symmetric equilibrium with trade and migration

In this section we briefly show that with free trade and migration, qualitatively similar restrictions on the parameter space and the shape of $f[\cdot]$ ensure existence and uniqueness of the symmetric equilibrium. In analogy to Section A.3 we show that qualitatively similar conditions are needed for quasiconcavity of the profit function and existence of an interior solution. Log diagonal dominance is shown to hold at the equilibrium point, guaranteeing uniqueness of the second stage wage equilibrium. Then, we describe conditions under which the symmetric alternating pattern is the only pattern consistent with free entry when migration cost are non-prohibitive.

Quasi-concavity of profits. Firm i 's labor supply function with integrated labor markets in the general case (asymmetric location pattern and domestic or foreign identity of neighbors) is given by

$$L_i^M = \sum_{c=\ell,r} (L_{i,c}^n + L_{i,c}^m) \quad \text{with} \quad (\text{A.60})$$

$$L_{i,c}^n = \begin{cases} \frac{L}{H} F[d_{i,c}^n] & \text{if } d_{i,c}^n > -d_{i,c'}^n \\ 0 & \text{otherwise} \end{cases} \quad \text{and} \quad L_{i,c}^m = \begin{cases} \frac{L}{H} (1-\lambda) F[d_{i,c}^m] & \text{if } d_{i,c}^m > -d_{i,c'}^m \\ 0 & \text{otherwise} \end{cases}$$

for $c, c' = \ell, r, c \neq c'$. We now denote with c, c' the *relevant* competitor of firm i on either side.⁵⁰ Note that with integrated labor markets and positive migration cost, firm i 's relevant competitor for natives on a given side may be a different firm than firm i 's relevant competitor migrant labor *on that same side*.⁵¹

The cutoff for native workers on side $c = \ell, r$, $d_{i,c}^n$, is determined by $w_i f[d_{i,c}^n] = w_c f[m_c^n - d_{i,c}^n]$ if c is a domestic firm and by $w_i f[d_{i,c}^m] = w_c^* f[m_c^n - d_{i,c}^m](1-\lambda)$ if c is a foreign firm. The cutoff for native workers, $d_{i,c}^m$, is determined by $w_i f[d_{i,c}^m] = w_c f[m_c^m - d_{i,c}^m]$ if c is a domestic firm and by $w_i f[d_{i,c}^m](1-\lambda) = w_c^* f[m_c^m - d_{i,c}^m]$ if c is a foreign firm. We denote with m_c^n, m_c^m the distance to the respective relevant competitor. The slope of a firm's supply of native

⁵⁰ The relevant competitor can be identified in similar way as explained in footnote ¹⁶.

⁵¹ With asymmetric locations and positive λ , it is conceivable that the competitor for natives, firm $i+1$, is overbid by a foreign firm $i+2$ with regard to migrants but not natives.

labor when competing with a firm in the other country is

$$\frac{\partial L_{i,c}^n}{\partial w_i} = \frac{L}{H} f[d_{i,c}^n] \frac{dd_{i,c}^n}{dw_i} = \frac{L}{H} \frac{f[d_{i,c}^n]^2}{-w_i f'[d_{i,c}^n] - w_c^*(1-\lambda) f'[m_{i,c}^n - d_{i,c}^n]} \quad (\text{A.61})$$

and when competing with a firm from the same country it is

$$\frac{\partial L_{i,c}^n}{\partial w_i} = \frac{L}{H} f[d_{i,c}^n] \frac{dd_{i,c}^n}{dw_i} = \frac{L}{H} \frac{f[d_{i,c}^n]^2}{-w_i f'[d_{i,c}^n] - w_c f'[m_{i,c}^n - d_{i,c}^n]}. \quad (\text{A.62})$$

Analogously, the slope of the supply of migrant labor when the competitor is foreign is

$$\frac{\partial L_{i,c}^m}{\partial w_i} = \frac{L}{H} (1-\lambda) f[d_{i,c}^m] \frac{dd_{i,c}^m}{dw_i} = \frac{L}{H} \frac{(1-\lambda) f[d_{i,c}^m]^2}{-w_i f'[d_{i,c}^m] - \frac{w_c^*}{(1-\lambda)} f'[m_{i,c}^m - d_{i,c}^m]}. \quad (\text{A.63})$$

When the competitor is in the same country, it is

$$\frac{\partial L_{i,c}^m}{\partial w_i} = \frac{L}{H} (1-\lambda) f[d_{i,c}^m] \frac{dd_{i,c}^m}{dw_i} = \frac{L}{H} \frac{(1-\lambda) f[d_{i,c}^m]^2}{-w_i f'[d_{i,c}^m] - w_c f'[m_{i,c}^m - d_{i,c}^m]}. \quad (\text{A.64})$$

As above, quasiconcavity of profits holds if condition (A.8) is fulfilled, and if the labor supply function becomes flatter at the kinks. Using the defining equations for the cutoffs with a foreign neighbor and Equations (A.61)-(A.64) it is straightforward to show that all possible cases (the competitor who is overbid is foreign or domestic, the next competitor is foreign or domestic), the respective labor supply schedule for natives and migrants becomes flatter at the kinks. Hence, quasiconcavity obtains under the restriction that the elasticity of marginal labor supply is not too large if positive. A similar condition on the choke price as above ensures that all solutions are interior.

Diagonal dominance at the equilibrium point. Using

$$\frac{\partial^2 L_i}{\partial w_i^2} = \sum_{c=\ell, r} \left(\frac{\partial^2 L_{i,c}^n}{\partial w_i^2} + \frac{\partial^2 L_{i,c}^m}{\partial w_i^2} \right) \quad \text{and} \quad \frac{\partial^2 L_i}{\partial w_i \partial w_c} = \frac{\partial^2 L_{i,c}^n}{\partial w_i \partial w_c} + \frac{\partial^2 L_{i,c}^m}{\partial w_i \partial w_c} \quad (\text{A.65})$$

where, if c is a foreign competitor,

$$\begin{aligned}
\frac{\partial^2 L_{i,c}^n}{\partial w_i^2} &= \frac{L}{2H} \left(3f'[d_{i,c}^n] \left(\frac{dd_{i,c}^n}{dw_i} \right)^2 + (w_i f''[d_{i,c}^n] - w_c^* (1-\lambda) f''[m_{i,c}^n - d_{i,c}^n]) \left(\frac{dd_{i,c}^n}{dw_i} \right)^3 \right) \\
\frac{\partial^2 L_{i,c}^m}{\partial w_i^2} &= \frac{L}{2H} \left(3(1-\lambda) f'[d_{i,c}^m] \left(\frac{dd_{i,c}^m}{dw_i} \right)^2 + (w_i (1-\lambda) f''[d_{i,c}^m] - w_c^* f''[m_{i,c}^m - d_{i,c}^m]) \left(\frac{dd_{i,c}^m}{dw_i} \right)^3 \right) \\
\frac{\partial^2 L_{i,c}^n}{\partial w_i \partial w_c} &= \frac{L}{2H} \left(2f'[d_{i,c}^n] \frac{dd_{i,c}^n}{dw_i} \frac{dd_{i,c}^n}{dw_c^*} + (1-\lambda) f'[m_{i,c}^n - d_{i,c}^n] \left(\frac{dd_{i,c}^n}{dw_i} \right)^2 \right. \\
&\quad \left. + (w_i f''[d_{i,c}^n] - w_c^* (1-\lambda) f''[m_{i,c}^n - d_{i,c}^n]) \left(\frac{dd_{i,c}^n}{dw_i} \right)^2 \frac{dd_{i,c}^n}{dw_c^*} \right) \\
\frac{\partial^2 L_{i,c}^m}{\partial w_i \partial w_c} &= \frac{L}{2H} \left(2(1-\lambda) f'[d_{i,c}^m] \frac{dd_{i,c}^m}{dw_i} \frac{dd_{i,c}^m}{dw_c^*} + f'[m_{i,c}^m - d_{i,c}^m] \left(\frac{dd_{i,c}^m}{dw_i} \right)^2 \right. \\
&\quad \left. + (w_i (1-\lambda) f''[d_{i,c}^m] - w_c^* f''[m_{i,c}^m - d_{i,c}^m]) \left(\frac{dd_{i,c}^m}{dw_i} \right)^2 \frac{dd_{i,c}^m}{dw_c^*} \right)
\end{aligned}$$

and

$$\begin{aligned}
\frac{\partial^2 L_{i,c}^n}{\partial w_i^2} &= \frac{L}{2H} \left(3f'[d_{i,c}^n] \left(\frac{dd_{i,c}^n}{dw_i} \right)^2 + (w_i f''[d_{i,c}^n] - w_c f''[m_{i,c}^n - d_{i,c}^n]) \left(\frac{dd_{i,c}^n}{dw_i} \right)^3 \right) \\
\frac{\partial^2 L_{i,c}^m}{\partial w_i^2} &= \frac{L}{2H} \left(3(1-\lambda) f'[d_{i,c}^m] \left(\frac{dd_{i,c}^m}{dw_i} \right)^2 + (w_i f''[d_{i,c}^m] - w_c f''[m_{i,c}^m - d_{i,c}^m]) \left(\frac{dd_{i,c}^m}{dw_i} \right)^3 \right) \\
\frac{\partial^2 L_{i,c}^n}{\partial w_i \partial w_c} &= \frac{L}{2H} \left(2f'[d_{i,c}^n] \frac{dd_{i,c}^n}{dw_i} \frac{dd_{i,c}^n}{dw_c} + f'[m_{i,c}^n - d_{i,c}^n] \left(\frac{dd_{i,c}^n}{dw_i} \right)^2 \right. \\
&\quad \left. + (w_i f''[d_{i,c}^n] - w_c f''[m_{i,c}^n - d_{i,c}^n]) \left(\frac{dd_{i,c}^n}{dw_i} \right)^2 \frac{dd_{i,c}^n}{dw_c} \right) \\
\frac{\partial^2 L_{i,c}^m}{\partial w_i \partial w_c} &= \frac{L}{2H} \left(2(1-\lambda) f'[d_{i,c}^m] \frac{dd_{i,c}^m}{dw_i} \frac{dd_{i,c}^m}{dw_c} + f'[m_{i,c}^m - d_{i,c}^m] \left(\frac{dd_{i,c}^m}{dw_i} \right)^2 \right. \\
&\quad \left. + (w_i f''[d_{i,c}^m] - w_c^* f''[m_{i,c}^m - d_{i,c}^m]) \left(\frac{dd_{i,c}^m}{dw_i} \right)^2 \frac{dd_{i,c}^m}{dw_c} \right)
\end{aligned}$$

if c is a domestic competitor, we can show that Equation (A.15) also holds for the case of migration. It follows that log diagonal dominance in accordance with Equation (A.11) also holds. Hence, the second-stage wage equilibrium with migration is unique.

Existence and uniqueness of the symmetric alternating location equilibrium.

A condition on the fixed cost relative to the size of the labor force similar to (A.17) can be derived that ensures existence of a symmetric second-stage equilibrium with symmetric distance pattern and zero profits. Moreover, it holds that $G^M[m] := \rho^T[m, \lambda]\psi^T[m] - g^M[m, \lambda]$ is monotonously increasing in m , hence the symmetric zero-profit solution is unique and second-stage profits for symmetric distance vectors are decreasing in N . To show that under the same assumption on consistency of beliefs as described in Section 2.2, the symmetric alternating distance pattern is the unique equilibrium as defined in (15), we need again a restriction on the magnitude of the change in the wage markup relative to the change in $g^M[m, \lambda]$. Analogously to the proof in Section A.3, a small enough level of β always assures that this condition holds. By the same logic as outlined in Section A.3, uniqueness of the symmetric equilibrium can be proven by showing that moving from the symmetric alternating equilibrium to any asymmetric pattern with the same or a larger number of firms implies negative profits for at least on firm.

Note that besides the alternating pattern another fully symmetric loction structure is conceivable, namely, one where each firm has one domestic neighbor and one foreign neighbor. However, as we show next, the alternating pattern is the one that maximizes labor supply per firm. With equal wages and one relevant domestic neighbor at a distance $m_{i,c}$, labor supply from the side where the domestic neighbor is located is given by

$$L_i^{M,D} = \frac{L}{2H}(2 - \lambda) \int_0^{m_{i,c}/2} f[d]dd \quad (\text{A.66})$$

for $\lambda \in [0, \bar{\lambda}]$. If, instead, the competitor at distance $m_{i,c}$ is foreign, the labor supply is

$$L^{M,F} = \frac{L}{2H} \int_0^{d_{i,c}^n} f[d]dd + \frac{L}{2H}(1 - \lambda) \int_0^{d_{i,c}^m} f[d]dd \quad (\text{A.67})$$

where $d_{i,c}^m \leq \frac{m_{i,c}}{2}$, $\frac{m_{i,c}}{2} \leq d_{i,c}^n < m$ and $d_{i,c}^m + d_{i,c}^n = m_{i,c}$. For notational convenience I henceforth set $L/(2H) = 1$. Then, the difference in supply of efficiency units for a given wage results as

$$L^{M,F} - L^{M,D} = \int_{m_{i,c}/2}^{d_{i,c}^n} f[d]dd - (1 - \lambda) \int_{d_{i,c}^m}^{m_{i,c}/2} f[d]dd. \quad (\text{A.68})$$

Using the fact that with symmetric wages $\int_{d_{i,c}^m}^{m_{i,c}/2} f[d]dd = \int_{m_{i,c}/2}^{d_{i,c}^n} f[m_{i,c} - d]dd$ this can be rewritten as

$$L^{M,F} - L^{M,D} = \int_{m_{i,c}/2}^{d_{i,c}^n} (f[d] - (1 - \lambda)f[m_{i,c} - d])dd \geq 0. \quad (\text{A.69})$$

The inequality follows from $f[d] - (1 - \lambda)f[m_{i,c} - d] \geq 0 \forall \frac{m_{i,c}}{2} \leq d_{i,c}^n < m_{i,c}$. Hence, in the symmetric equilibrium the labor supply for a given wage is (weakly) larger if the neighbor is foreign. If $\lambda = 0$, labor supply is identical in both cases. Hence, by the same logic that rules out asymmetric distance patterns with a number of firms larger or equal to the number of firms in the symmetric alternating zero-profit solution, non-alternating symmetric distance patterns cannot constitute an equilibrium as defined in (15), unless migration cost are zero. In the last case, the symmetric alternating and non-alternating equilibrium are indistinguishable.

A.7.2 The elasticity of labor supply

The elasticity of labor supply in the symmetric alternating equilibrium is defined as $\frac{\partial L^{ES,M}}{\partial w_i} \frac{w_i}{L^{ES,M}}$. From (37), (35), and (36), we obtain

$$\frac{\partial L^{ES,M}}{\partial w_i} = \frac{L}{H} \frac{dd_i^n}{dw_i} f[d_i^n] + (1 - \lambda) \frac{L}{H} \frac{dd_i^m}{dw_i} f[d_i^m] \quad \text{with} \quad (\text{A.70})$$

$$\frac{dd_i^n}{dw_i} = \frac{f[d_i^n]}{-w_i f'[d_i^n] - w^*(1 - \lambda) f'[m - d_i^n]} \quad (\text{A.71})$$

$$\frac{dd_i^m}{dw_i} = \frac{(1 - \lambda) f[d_i^m]}{-w_i(1 - \lambda) f'[d_i^m] - w^* f'[m - d_i^m]}. \quad (\text{A.72})$$

Evaluating $\frac{\partial L^{ES,M}}{\partial w_i} \frac{w_i}{L^{ES,M}}$ at the symmetric equilibrium, where it holds that $w_i = w^* \equiv 1$, $d_i^n = d^n$, $d_i^m = d^m = m - d^n$ and $f[d^n] = (1 - \lambda)f[d^m]$, we obtain

$$\begin{aligned} \eta^M &= \left. \frac{\partial L^{ES,M}}{\partial w_i} \frac{w_i}{L^{ES}} \right|_{w_i=w} = \frac{L}{H} \left(\frac{f[d^n]^2}{-f'[d^n] - (1 - \lambda)f'[m - d^n]} + \frac{(1 - \lambda)^2 f[d^m]^2}{-(1 - \lambda)f'[d^m] - f'[m - d^m]} \right) \\ &\quad \times \frac{1}{\frac{L}{H} \left(\int_0^{d^n} f[d] dd + (1 - \lambda) \int_0^{d^m} f[d] dd \right)} \\ &= \frac{2f[d^n]^2}{f'[d^n] + (1 - \lambda)f'[d^m]} \cdot \frac{-1}{\int_0^{d^n} f[d] dd + (1 - \lambda) \int_0^{d^m} f[d] dd} \end{aligned} \quad (\text{A.73})$$

as displayed in (39). The elasticity of labor supply decreases in m :

$$\eta_m^M = \eta^M \left[\frac{2f'[d^n]}{f[d^n]} \frac{dd^n}{dm} - \frac{-f''[d^n] \frac{dd^n}{dm} - (1 - \lambda) f''[d^m] \frac{dd^m}{dm}}{-f'[d^n] - (1 - \lambda)f'[d^m]} - \frac{f[d^n]}{m\theta^M} \right] < 0, \quad (\text{A.74})$$

where $\frac{dd^n}{dm} = \frac{(1-\lambda)f'[d^m]}{f'[d^n] + (1-\lambda)f'[d^m]} > 0$ and $\frac{dd^m}{dm} = \frac{f'[d^n]}{f'[d^n] + (1-\lambda)f'[d^m]} > 0$. Furthermore, provided that $f'''[\cdot]$ is not too positive, η^M decreases in λ :

$$\eta_\lambda^M = \eta^M \left[\underbrace{\frac{2f'[d^n]}{f[d^n]} \frac{dd^n}{d\lambda}}_{<0} + \underbrace{\frac{f''[d^n] \frac{dd^n}{d\lambda} + (1-\lambda)f''[d^m] \frac{dd^m}{d\lambda} + f'[d^m]}{-f'[d^n] - (1-\lambda)f'[d^m]}}_{<0} + \underbrace{\frac{F[d^m]}{F[d^n] + (1-\lambda)F[d^m]}}_{>0} \right] < 0 \quad (\text{A.75})$$

with $\frac{dd^n}{d\lambda} = \frac{f[d^m]}{-f'[d^n] - (1-\lambda)f'[d^m]} > 0$ and $\frac{dd^m}{d\lambda} = -\frac{dd^n}{d\lambda} < 0$. $\eta_\lambda^M < 0$ follows from the fact that the first term in the brackets (in absolute terms) exceeds the third, since

$$\frac{2f'[d^n]}{f[d^n]} \frac{dd^n}{d\lambda} = 2 \underbrace{\frac{f[d^m]}{f[d^n]} \cdot \frac{f'[d^n]}{f'[d^n] + (1-\lambda)f'[d^m]}}_{\geq 1} \geq \underbrace{\frac{F[d^m]}{F[d^n] + (1-\lambda)F[d^m]}}_{< 1}. \quad (\text{A.76})$$

A.7.3 Analytical details of the proof of Proposition 3

The number of firms is too large in the migration equilibrium. The social planner solves the same maximization problem as in Appendix A.5, additionally taking into account the integrated labor market.⁵² The first order condition of the planner then obtains as

$$\frac{Lf[d^n]}{L\theta^M - \frac{\alpha H}{m}} = 1 + \frac{m}{4\gamma H}. \quad (\text{A.77})$$

where d^n, θ^M are shorthands for $d^n[m, \lambda], \theta^M[m, \lambda]$, respectively. A comparison with the market solution (42) shows that, as before, the number of firms in the market equilibrium is too large if the markup distortion is larger than the productivity distortion. We can show that this is the case in the migration equilibrium with non-prohibitive λ . The relevant condition is $\psi^M > \frac{\theta^M}{f[d^n]}$. Inserting for ψ^M this is equivalent to $1 - \frac{m\theta^M(f'[d^n] + (1-\lambda)f'[d^m])}{2f[d^n]^2} > \frac{\theta^M}{f[d^n]}$. This, in turn, holds if $1 - \frac{m\theta^M f'[d^n]}{2f[d^n]^2} > \frac{\theta^M}{f[d^n]}$, since $-f'[d^m](1-\lambda)/(2f[d^n]^2) \geq 0$. Rewriting the condition leads to $f[d^n] > \theta^M + \frac{m}{2} \frac{f'[d^n]}{f[d^n]} \theta^M$. We will show below that $f[\frac{d^n}{2}] \geq \theta^M$. Then, this inequality holds if

$$f[d^n] > f\left[\frac{d^n}{2}\right] + \frac{m}{2} \frac{f'[d^n]}{f[d^n]} \theta^M. \quad (\text{A.78})$$

⁵² Note that this assumes that either the planner maximizes welfare for both countries or takes as given that a planner in the foreign country solves the exact same problem.

Concavity of $f[\cdot]$ implies that $f[d^n] \geq f\left[\frac{d^n}{2}\right] + f'[d^n]\frac{d^n}{2}$. Moreover, we have that $f\left[\frac{d^n}{2}\right] + f'[d^n]\frac{d^n}{2} > f\left[\frac{d^n}{2}\right] + f'[d^n]\frac{m}{2}\frac{\theta^M}{f[d^n]}$ because $m \geq d^n$ and $\theta^M > f[d^n]$. Therefore, (A.78) holds a fortiori. Hence, the markup distortion exceeds the productivity distortion and consequently, the number of firms in the market equilibrium with migration is too large.⁵³

Proof that $\theta^M \leq f\left[\frac{d^n}{2}\right]$. Using the expression for θ^M in (38) and Jensen's inequality which states that $f[E[x]] \geq E[f[x]]$ for concave functions $f[x]$, we can state

$$\theta^M = \frac{1}{m} \int_0^{d^n} f[d]dd + (1 - \lambda) \frac{1}{m} \int_0^{d^m} f[d]dd \leq \frac{d^n}{m} f\left[\frac{d^n}{2}\right] + (1 - \lambda) \frac{d^m}{m} f\left[\frac{d^m}{2}\right] \quad (\text{A.79})$$

Since $d^n + d^m = m$, we have that $\theta^M \leq \frac{d^n}{m} f\left[\frac{d^n}{2}\right] + (1 - \lambda) \frac{d^m}{m} f\left[\frac{d^m}{2}\right]$. This reduces to $\theta^M \leq f\left[\frac{d^n}{2}\right]$ provided that $(1 - \lambda) f\left[\frac{d^m}{2}\right] \leq f\left[\frac{d^n}{2}\right]$. From (35) and (36) it follows that a symmetric equilibrium is characterized by $(1 - \lambda) = f[d^n]/f[d^m]$, so the condition becomes $\frac{f\left[\frac{d^m}{2}\right]}{f\left[\frac{d^n}{2}\right]} \leq \frac{f[d^m]}{f[d^n]}$, which is implied by $d^m \leq d^n$ and $f''[\cdot] \leq 0$. This completes the proof.

⁵³ There is a subtle point to this proof in that $\theta^M[m, \lambda]$ is not necessarily concave in m , if there is migration. As a result, the social welfare function is not globally concave. However, it can be shown that the first order condition in A.77 still describes a global maximum and that the social welfare function is monotonously increasing in the relevant range. Details of the proof are available upon request.

A.7.4 Proof of Proposition 4

Totally differentiating the equilibrium condition (42) yields $\hat{m} = C \cdot \hat{\lambda}$ where C is given by⁵⁴

$$C = \frac{g_\lambda^M - \rho^T \psi_\lambda^M}{-g_m^M + \rho^T \psi_m^M + \psi_m^M \rho_m^T} \frac{\lambda}{m} \leq 0 \quad \text{with} \quad (A.80)$$

$$g_\lambda^M = \frac{L\theta_\lambda^M}{L\theta^M - \alpha N} - \frac{L\theta^M}{(L\theta^M - \alpha N)^2} L\theta_\lambda^M > 0 \quad \text{and} \quad \theta_\lambda^M = -\frac{1}{m} \int_0^{d^m} f[d] dd < 0 \quad (A.81)$$

$$g_m^M = \frac{L\theta_m^M}{L\theta^M - \alpha N} - \frac{L\theta^M}{(L\theta^M - \alpha N)^2} \left(L\theta_m^M + \frac{\alpha N}{m} \right) < 0 \quad \text{and} \quad \theta_m^M = \frac{1}{m} (f[d^m] - \theta^M) < 0 \quad (A.82)$$

$$\psi_\lambda^M = -\frac{1}{(\eta_\lambda^M)^2} \cdot \eta_\lambda^M > 0 \quad \text{with } \eta_\lambda^M \text{ as in (A.75)} \quad (A.83)$$

$$\psi_m^M = -\frac{1}{(\eta_m^M)^2} \cdot \eta_m^M > 0 \quad \text{with } \eta_m^M \text{ as in (A.74)} \quad (A.84)$$

$$\rho_m^T = \frac{1}{2\gamma H} > 0. \quad (A.85)$$

While the denominator of C is always positive (a larger firm size m decreases the markup needed for zero profits g^M and increases both the price markup and the wage markup), the sign of the numerator depends on whether the effect of λ on g^M (which is positive) is stronger than the effect on the wage markup (which is also positive). In either case, prices fall as migration costs fall.

The effect on average income is ambiguous. While the partial effect of lower migration costs is positive, there is a countervailing effect when the general equilibrium adjustments lead to firm exit. In either case, however, real income increases when migration costs fall, as the decrease in prices overcompensates the potential decrease in average income. We show this by log-differentiating real income $\frac{\theta^M}{p} = \frac{L\theta^M - \frac{\alpha H}{m}}{L}$ as obtained by rewriting (42):

$$d \ln \left[\frac{\theta^M}{p} \right] = \frac{\partial \ln \left[\frac{\theta^M}{p} \right]}{\partial \lambda} \lambda \cdot \hat{\lambda} + \frac{\partial \ln \left[\frac{\theta^M}{p} \right]}{\partial m} m \cdot \hat{m} \quad (A.86)$$

⁵⁴Note that for notational convenience here and in the following we omit the functional dependence of $g^M, \psi^M, \rho^M, \theta^M, d^n$ on m and, where relevant, on λ .

with

$$\frac{\partial \ln \left[\frac{\theta^M}{p} \right]}{\partial \lambda} = \frac{L\theta_\lambda^M}{L\theta^M - \frac{\alpha H}{m}} < 0 \quad \text{and} \quad (\text{A.87})$$

$$\frac{\partial \ln \left[\frac{\theta^M}{p} \right]}{\partial m} = \frac{L\theta_m^M + \frac{\alpha H}{m^2}}{L\theta^M - \frac{\alpha H}{m}} > 0. \quad (\text{A.88})$$

In these equations $\theta_\lambda^M = -\frac{1}{m} \int_0^{d^m} f[d]dd < 0$ and $\theta_m^M = \frac{1}{m} (f[d^n] - \theta^M) < 0$. Note that $\frac{\partial \ln \left[\frac{\theta^M}{p} \right]}{\partial m} > 0$ in the relevant range follows from (A.77). Hence, the log-change in real income induced by a decrease in λ is clearly positive, if \hat{m} is also positive. To show that real income also increases if \hat{m} is negative we use (A.80) as well as (A.87) and (A.88) to rewrite (A.86) as

$$\begin{aligned} d \ln \left[\frac{\theta^M}{p} \right] &= \frac{\lambda}{(L\theta^M - \alpha N)(-g_m^M + \rho\psi_m^M + \psi^M\rho_m)} \\ &\times \left[\left(L\theta_m^M + \frac{\alpha N}{m} \right) (g_\lambda^M - \rho\psi_\lambda^M) + (-g_m^M + \rho\psi_m^M + \psi^M\rho_m) L\theta_\lambda^M \right] \hat{\lambda}. \end{aligned} \quad (\text{A.89})$$

We know that the first fraction on the right-hand side above is positive, hence we must show that the square-bracketed term is negative. Using

$$\left(L\theta_m^M + \frac{\alpha N}{m} \right) g_\lambda^M = \left[\frac{L\theta_m^M + \frac{\alpha N}{m}}{L\theta^M - \alpha N} - \frac{L\theta^M (L\theta_m^M + \frac{\alpha N}{m})}{(L\theta^M - \alpha N)^2} \right] \cdot L\theta_\lambda^M \quad (\text{A.90})$$

and

$$L\theta_\lambda^M g_m^M = \left[\frac{L\theta_m^M}{L\theta^M - \alpha N} - \frac{L\theta^M (L\theta_m^M + \frac{\alpha N}{m})}{(L\theta^M - \alpha N)^2} \right] \cdot L\theta_\lambda^M \quad (\text{A.91})$$

we can reduce the expression in squared brackets on the right-hand side of (A.89) to

$$L\theta_\lambda^M \left(\frac{\frac{\alpha N}{m}}{L\theta^M - \alpha N} + \psi_m^M \rho + \rho_m \psi^M \right) - \left(L\theta_m^M + \frac{\alpha N}{m} \right) \rho \psi_\lambda^M. \quad (\text{A.92})$$

This is negative since $\theta_\lambda^M < 0$ and $\psi_\lambda^M > 0$. Hence, a decrease in λ raises real income also if it leads to exit of firms. This completes the proof of proposition 4.

A.7.5 Robustness with respect to the specification of migration costs

The proofs of proposition 3 and 4 reveal that our results are valid for more general specifications of migration costs. The positive welfare effect of the potential of migration established in proposition 3 stems from a first-order welfare gain due the reduction of the markup distortion. Hence, the validity of proposition 3 is maintained, provided that the excess-entry property of the autarky equilibrium is preserved. The proof of proposition 4 shows that positive welfare gains from lower migration costs occur, provided that $\theta_\lambda^M < 0$ and $\eta_\lambda^M < 0$, and that the excess-entry result holds. It is relatively straightforward that this holds for a wide range of migration costs specifications.