



Urban agglomeration and heterogeneous firms: a synthesis of Helpman and Melitz

Yiming Zhou¹

Received: 5 October 2019 / Accepted: 11 April 2020 / Published online: 7 May 2020
© Springer-Verlag GmbH Austria, part of Springer Nature 2020

Abstract

This study revisits the seminal findings by Helpman in 1998 in a more comprehensive setup with Melitz-type firm heterogeneity. Our model features how a change in firm heterogeneity affects agglomeration in an urban-space economy. We contribute to the literature by analytically giving explicit solutions for the threshold values of the housing preference and transport costs that are crucial in determining the equilibrium. We also obtain new findings in the case of asymmetric housing stocks. Our results provide alternative explanations to stylized facts such as “ghost cities” in countries undertaking high-speed urbanization and real estate development.

Keywords Urban agglomeration · Housing congestion · Firm heterogeneity

JEL Classification F12 · F22 · R12

1 Introduction

In the past several decades, the new economic geography (NEG) literature has contributed to our understanding of the mechanism of economic agglomeration across space. However, as a pioneering work of this literature, the core-periphery model (Krugman 1991a) has also been criticized because it does not fit well contemporary space-economies (Murata and Thisse 2005). Nowadays, in a growing number of countries, a few large metropolitan areas produce a growing share of their gross domestic product (GDP) and, consequently, urban costs take an

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s00712-020-00698-5>) contains supplementary material, which is available to authorized users.

✉ Yiming Zhou
zhouchn@hit.edu.cn

¹ School of Management, Harbin Institute of Technology, Harbin 150001, China

increasing amount of household expenditures (Proost and Thisse 2019). As a result, the main spatial dispersion force seems to stem from urban costs, instead of the immobile expenditure in the agricultural sector assumed in the core-periphery model.

In this regard, a notable and highly cited paper is Helpman (1998) who introduced a housing market into an economic geography model in which all workers are mobile. In his model, the major centrifugal force stems from housing congestion. Contrary to the results of the core-periphery model, Helpman (1998) finds that lower transport costs lead to less agglomeration while higher transport costs may lead to larger agglomeration. He shows how interplay between the centripetal force originating from the differentiated goods sector and the centrifugal force stemming from the housing market reshapes the spatial configuration of an urban-space economy. This seminal work has inspired and stimulated many related studies (see, e.g., Murata and Thisse 2005 and Pflüger and Tabuchi 2010, to name a few) that further explore the role of urban elements in the functioning of spatial economies. More recently, Helpman's setup has also been widely employed in recent literature on quantitative spatial economics (see, e.g., Redding 2016 and Redding and Rossi-Hansberg 2017).

However, two shortcomings of Helpman (1998) have been neglected and received little attention in the literature. First, the main results of Helpman (1998) are derived through numerical simulations. Although his results indicate that the housing preference, the elasticity of substitution, and the transport costs interact and have significant impacts on spatial equilibrium, closed-form solutions for the threshold values of the above variables were not derived. Second, the firms are assumed to be homogeneous in productivity in Helpman's model. However, a large body of recent literature reveals that firm heterogeneity in productivity has significant impacts on trade, firm selection, and economic geography (see, e.g., Melitz 2003; Baldwin and Okubo 2006 and Okubo et al. 2010).

This study aims to overcome these two shortcomings by introducing firm heterogeneity in productivity *à la* Melitz (2003) into the model of Helpman (1998). By doing so, the contributions of this study are threefold. First, we analytically give explicit solutions for the threshold values of the housing preference and the transport costs that are crucial in determining the spatial configuration. Second, we analytically show that, besides the elasticity of substitution across varieties, the degree of firm heterogeneity enters the threshold values as well. Our model features how a change in firm heterogeneity affects the spatial equilibrium. By comparing our result with that of related studies, we reveal that the firm heterogeneity affects spatial equilibrium in a manner very similar to trade liberalization. Third, we have new findings that are in contrast with Helpman (1998) if the housing stocks are unevenly distributed across regions. We explain intuitively how and why our results differ. In particular, our results provide alternative theoretical explanations for some stylized facts such as "ghost cities" in countries undertaking high-speed urbanization and real estate development.

Our findings are as follows: First, the symmetric equilibrium is always stable (unstable) if housing preference is higher (lower) than a threshold value. For the middle level of housing preference, the symmetric equilibrium is stable if

transport costs are lower than a threshold value (i.e., the break point). On the other hand, the full agglomeration is always stable (unstable) if the housing preference is lower (higher) than a threshold value. For the middle level of housing preference, the full agglomeration is stable if the transport costs are higher than a threshold value (i.e., the sustain point).

Second, a comparative static analysis shows that an increase in firm heterogeneity works in favor of dispersion. Intuitively, an increase in firm heterogeneity, among others, implies a larger proportion of high-productivity firms and a higher propensity to export. As a larger proportion of globally produced varieties becomes available and can be imported at lower prices, living in the larger market becomes less important, which attenuates the centripetal force stemming from the home-market effect.

Third, we find that the main findings in symmetric settings remain largely robust in the case of asymmetric housing stocks across regions. Our results show that the share of the population is *more than proportional* to its housing stock in the larger region. This is in contrast with Helpman's prediction that the share of the population is *proportional* to each region's housing stock when transport costs are negligible. The difference comes from the firm heterogeneity in productivity. In Helpman's model, firms are homogeneous and a representative resident has access to all brands of varieties if the transport costs are negligible. She migrates to the region with less housing congestion until each region's population becomes proportional to its housing stock.

In contrast, in our setup with firm heterogeneity, due to the existence of fixed export costs, only a proportion of high productivity firms export, even when the transport costs are negligible. This gives the consumers in the larger region better access to the differentiated goods, and this merit is just enough to compensate for the housing congestion resulting from a more than proportional share of population in that region. Moreover, we find that if the housing preference is low, it is very possible that the region with much more housing stock turns out to be the periphery region. This finding provides alternative theoretical explanations to the phenomenon of "ghost cities" in countries undertaking high-speed urbanization and real estate development.

Our study is related to the following literature. Tabuchi (1998) proposed a model which unifies economics *à la* Alonso (1964) and NEG in a single framework. His results depict a structural transition from dispersion to agglomeration, and then redispersion as transport costs fall. However, his analysis is limited to two extreme cases of zero and infinite transport costs. Murata and Thisse (2005) developed a more tractable framework with iceberg urban commuting costs that affect a worker's effective labor supply and her income. Contrary to the results of the core-periphery model, they show that mobile workers are unwilling to agglomerate in order to alleviate the urban costs, unless the transport costs are so high that the benefit of having all varieties locally produced outweighs the urban costs. The current study differs from these studies in that the consumption of housing and, thus, the dispersion force from fixed housing stock are not considered in their models. Moreover, firm heterogeneity in productivity is assumed away in their models.

Although the consumption of land and its use in production are considered in Pflüger and Tabuchi (2010), they assume homogeneous firms in their setup.

This paper is closely related with Ehrlich and Seidel (2013) who introduce Melitz-type firm heterogeneity into the standard core-periphery model. They disclose that an increase in firm heterogeneity works in favor of agglomeration. We differ from them in that urban cost and its impact on agglomeration are not taken into account in their model.

In this regard, our results are consistent with those of Zhou (2018), who incorporates Melitz-type firm heterogeneity into the model of Murata and Thisse (2005). He shows that higher urban costs or lower transport costs foster dispersion, and an increase in firm heterogeneity works in favor of dispersion. Our setup differs from his in that the urban costs in Zhou (2018) take the form of commuting cost and land rent which reduce labor's effective income, while the preference for housing consumption and the disutility from housing congestion are not taken into account. While these urban elements coexist in reality, the current study enriches our understanding of the roles of urban costs and firm heterogeneity in urban agglomeration. Second, while the analysis of Zhou (2018) is limited to symmetric settings, this paper examines the scenario with asymmetric settings. Third, by classifying our results with those of related studies, we reveal that firm heterogeneity affects spatial equilibrium in a manner similar to trade liberalization.

Our results also add to the so-called 'new' NEG literature.¹ In this vein of research, Baldwin and Okubo (2006) introduce firm heterogeneity *à la* Melitz (2003) into the footloose capital model proposed by Martin and Rogers (1995). They show that firm heterogeneity leads to a sorting of the most productive firms into the larger regions. However, based on the footloose capital model, where the mobile factor repatriates all its earnings to its region of origin, their approach does not exhibit circular causality as in the standard core-periphery model. Okubo (2009) further derives gradual agglomeration (rather than catastrophic agglomeration) by introducing intermediate input linkages into the model of Baldwin and Okubo (2006). Among others, Nocke (2006) finds that each entrepreneur in a large market is more efficient than any entrepreneur in a smaller market, because competition is endogenously more intense in larger markets. Our setup differs in that the workers here are homogeneous and as in Melitz (2003), each firm is associated with a particular labor input coefficient (i.e., marginal cost). In a linear model, Okubo et al. (2010) show that the more productive firms are selected into larger markets when trade costs fall, but the less productive firms also find it profitable to locate in the larger market. It thus leads to a bell-shaped relationship between trade liberalization and international productivity gap. Assuming two types of firm productivity, Saito (2015) examines the organization and location decisions of heterogeneous firms with multi-plant operations and their consequences for regional productivity. This paper differs from that strand of literature in that these studies do not take into account the centrifugal force stemming from housing congestion.

¹ See Ottaviano (2011) for a review of this stream of literature.

The remainder of this paper is organized as follows. Section 2 introduces the model setting. Section 3 examines the short-run equilibrium. Section 4 examines the spatial equilibrium. The last section summarizes our main results.

2 The model

2.1 The spatial economy

Consider an economy involving two regions ($r = 1, 2$), one manufacturing sector producing differentiated products with increasing returns to scale, as well as a fixed stock of housing (S_r). The economy is endowed with L homogeneous and mobile workers, each supplying one unit of labor inelastically. Let λ denote the fraction of workers residing in region 1 so that the mass of workers in regions 1 and 2 is given by $L_1 = L\lambda$ and $L_2 = L(1 - \lambda)$, respectively. Obviously, labor supply in a region is determined by its population. Workers in the more populated region are supplied with a better variety choice of locally produced goods while imported brands are costly to transport. On the other hand, a worker lives in the region where she works, and purchases local housing services and all brands of differentiated products. As a result, housing costs are higher in the more populated region. Workers tend to agglomerate in the larger region in order to get better access of differentiated products and, on the other hand, to disperse in order to alleviate housing congestion. In the following, for simplicity, we mainly describe the economy in region 1, as region 2 is almost symmetric.

2.2 Consumption

Preferences are the same across consumers and each consumer in region 1 maximizes a Cobb–Douglas utility function given by

$$u_1 = h_1^\beta c_1^{1-\beta} \quad (1)$$

with

$$c_1 = \left[\int_{i \in \Omega} \hat{d}_1(i)^{\frac{\sigma-1}{\sigma}} di \right]^{\frac{\sigma}{\sigma-1}},$$

where h_1 represents housing consumption and $\hat{d}_1(i)$ denotes the consumption level of variety i . Since there are two regions and manufactured variety is generally tradable, $\hat{d}_1(i)$ may be the consumption level of a local or an imported product. The parameter, Ω , denotes the total mass of varieties that are endogenously determined in our model. A representative consumer maximizes utility subject to her budget constraint, and the individual demand for variety i is derived as

$$\hat{d}_1(i) = \frac{p_1(i)^{-\sigma} E_1(1 - \beta)}{P_1^{1-\sigma} L\lambda}, \quad (2)$$

where $p_1(i)$ is the consumer price of variety i , $P_1 \equiv (\int_{i \in \Omega} p_1(i)^{1-\sigma} di)^{\frac{1}{1-\sigma}}$ denotes the price index in region 1 and E_1 is the aggregate expenditure of region 1.

Each consumer spends a fraction β of her income on housing and, therefore, the aggregate expenditure on housing is $(E_1 + E_2)\beta$. Also, aggregate expenditure equals aggregate income, which is composed of labor income $L\lambda w_1 + L(1 - \lambda)w_2$ (where w_r is the wage rate) and income from housing ownership $(E_1 + E_2)\beta$. Thus, the aggregate income from housing equals $[L\lambda w_1 + L(1 - \lambda)w_2]\beta / (1 - \beta)$. As in Helpman (1998), the housing stocks are equally owned by all residents. As a result, income from housing by residents of region 1 equals the fraction λ of the total income from housing. Therefore, the aggregate expenditure of region 1 is derived as

$$E_1 = L\lambda w_1 + \frac{\beta\lambda[L\lambda w_1 + L(1 - \lambda)w_2]}{1 - \beta}. \quad (3)$$

2.3 Technology and production

Each variety of the differentiated goods is produced by a single firm under increasing returns and monopolistic competition. We follow Melitz (2003) in assuming that firms differ in their labor productivity φ which is drawn from a commonly-known distribution function. Firms do not know their productivity ex ante, but have to incur an investment (like R&D) to obtain this information. We denote this entry cost in terms of labor, that is $f_e w_r$. Based on this knowledge, firms decide about producing or exiting the industry in case their productivity is too low to make profits. After that, for x_i units of output of variety i , each firm has a specific input requirement according to $x_i(\varphi) = l_i(\varphi)\varphi$, where $l_i(\varphi)$ denotes marginal labor input by firm of type φ . Please note that, as in Melitz (2003), firms are heterogeneous w.r.t. their productivity although workers do not differ in their skills. This can be rationalized by arguing that each firm possesses a specific technology determining labor productivity of each employee.

Moreover, firms choose which markets to serve. In order to serve the local market, each firm is required to invest f units of labor as a fixed input. This investment could take the form of, say, equipment purchase or any marketing activities that are independent from variable costs. A similar argument applies for the export market such that firms have to hire additional f_x units of labor to sell to consumers abroad. When a differentiated good is shipped across regions, transport costs *à la* Samuelson (1954) occur: $\tau > 1$ units of the variety must be sent from the origin for one unit to arrive at destination. Since the varieties are symmetric, in the following, we drop the “ i ” to simplify the expressions.

Under Dixit–Stiglitz preferences, firms maximize their profits by choosing the optimal prices. For local sales and exports, the consumer prices are respectively derived as

$$p_1 = \frac{\sigma w_1}{(\sigma - 1)\varphi}, \quad p_{1x} = \frac{\tau\sigma w_1}{(\sigma - 1)\varphi}.$$

Together with the demand function (2), the revenues and profits of a representative firm in region 1 from local and foreign markets are derived as

$$R_1(\varphi) = \frac{p_1(\varphi)^{1-\sigma}}{p_1^{1-\sigma}} E_1(1 - \beta), \quad R_{1x}(\varphi) = \frac{p_{1x}(\varphi)^{1-\sigma}}{p_2^{1-\sigma}} E_2(1 - \beta),$$

$$\pi_1(\varphi) = R_1(\varphi)/\sigma - f w_1, \quad \pi_{1x}(\varphi) = R_{1x}(\varphi)/\sigma - f_x w_1,$$

where R_1 and π_1 are revenue and profit from local market whereas R_{1x} and π_{1x} are those from foreign market. Note that firms with higher productivity (higher φ) charge lower prices, sell more and earn higher profits.

We follow the literature on heterogeneous firms in assuming Pareto distributed productivity levels. Hence, the cumulative distribution function reads $G(\varphi) = 1 - \varphi^{-k}$, where $k > 0$ denotes the shape parameter. To simplify notation, as in Ehrlich and Seidel (2013), we have normalized the scale parameter to unity without loss of generality. This means that $\varphi = 1$ is the lowest productivity a firm can draw. As noted by Ehrlich and Seidel (2013), the Pareto distribution offers the advantage that the shape parameter k is a straightforward measure for the heterogeneity of firms. The variance of the Pareto distribution $Var(\varphi) = k/[(k - 1)^2(k - 2)]$ is strictly decreasing in k for $k > 2$.² A high value of k implies that it becomes less likely to draw a high productivity level φ . In other words, there are only a few very productive firms and many low-productive ones. In the extreme case of $k = \infty$, all firms are clustered at the lower bound $\varphi = 1$. We thus refer to lower levels of the shape parameter k as a more heterogeneous distribution of productivity levels.

3 Short-run equilibrium

We summarize the timeline of the behavior of firms in our model as follows: For a given allocation of population, firms decide whether to enter the industry until their expected profits equalize the entry costs. Based on their productivity draw, firms start producing as long as their profits are not negative and pay the market wage rate w_r . This is true for all firms with a productivity level φ that exceeds the cutoff level φ^* . Moreover, a subset of these domestically active firms with higher productivity may find it profitable to export to the foreign region. We refer to this situation as the short-run equilibrium without labor mobility. In the long-run, workers migrate across regions in search of the highest utility. If migrating to the other region promises higher real remuneration, the allocation of workers is modified such that firm entry and exit adjusts to meet the equilibrium condition of zero expected

² Assuming $k > 2$ is necessary for ensuring the Pareto distribution has finite variance. See also Helpman et al. (2004). Meanwhile, as in Ehrlich and Seidel (2013), we impose $k > \sigma - 1$ to ensure that the integrals of the average productivity of the Pareto distribution converge.

profits. The migration process terminates if either utilities are equalized across regions or all workers agglomerate in one jurisdiction.

In the short-run, for a given distribution of population, we first derive the local cutoff productivity level φ^* . To obtain φ^* , we combine the *free-entry condition* with the *zero-cutoff-profit condition*. Firms enter the industry as long as expected profits (from both local sales and exports) are sufficient to cover the fixed market entry costs. Formally, this free-entry condition is given by

$$(\varphi_1^*)^{-k} \bar{\pi}_1 = f_e w_1, \tag{4}$$

where $\bar{\pi}_1$ denotes average profits of surviving firms. Multiplied by the probability of surviving in competition, that is $1 - G(\varphi_1^*) = (\varphi_1^*)^{-k}$, we obtain expected profits before firm-specific productivity levels have been realized.

On the other hand, denoting by $\tilde{\varphi}_1$ and $\tilde{\varphi}_{1x}$ as the productivity levels of the average domestic and exporting firm, respectively, surviving firms can expect to earn $\pi_1(\tilde{\varphi}_1)$ domestically and $(\varphi_1^*/\varphi_{1x}^*)^k \pi_{1x}(\tilde{\varphi}_{1x})$ from exports. The term, $(\varphi_1^*/\varphi_{1x}^*)^k$, reflects the probability of becoming an exporter conditional on being active in the domestic market, with φ_{1x}^* denoting the export productivity cutoff. Firms will only start producing for domestic and export market as long as their revenues from the respective market cover the market-specific fixed costs. As a result, the marginal domestic and exporting firm will be formally given by

$$R_1(\varphi_1^*) = \sigma f w_1, \quad \text{and} \quad R_{1x}(\varphi_{1x}^*) = \sigma f_x w_1.$$

These two conditions can be used together with $R_{2x}(\varphi_{2x}^*) = R_1(\varphi^*) p_{2x}(\varphi_{2x}^*)^{1-\sigma} / p_1(\varphi^*)^{1-\sigma} = \sigma f_x w_2$ to establish a link between the domestic cutoff in region 1 and the exporter cutoff in region 2:

$$\varphi_{2x}^* = \tau(f_x/f)^{1/(\sigma-1)} (w_2/w_1)^{\sigma/(\sigma-1)} \varphi_1^*. \tag{5}$$

As in the literature, we assume $f_x > f$, which implies the reality that the domestic sales are generally more profitable than exporting. Please note that it is a common assumption in the literature to avoid the case where firms export without serving local consumers.³ Based on these insights, it is evident that the conditional export probability is limited to the range between zero and unity. Intuitively, a lower level of the shape parameter k (more heterogeneous in productivity) implies a higher export probability. By using Eq. (5), we can formulate the conditional export probability as

$$\left(\frac{\varphi_1^*}{\varphi_{1x}^*}\right)^k = \phi \left(\frac{f_x}{f}\right)^{\frac{k}{1-\sigma}} \left(\frac{w_1}{w_2}\right)^{\frac{\sigma k}{1-\sigma}} \left(\frac{\varphi_1^*}{\varphi_2^*}\right)^k, \tag{6}$$

where $\phi \equiv \tau^{-k} \in (0, 1)$.

³ Note that for symmetric regions, we have $w_1 = w_2$ and $\varphi_1^* = \varphi_2^*$ such that $\varphi_{2x}^* > \varphi_1^*$ implies $\varphi_{2x}^* > \varphi_2^*$ whereas $\varphi_{1x}^* > \varphi_2^*$ implies $\varphi_{1x}^* > \varphi_1^*$. For asymmetric regions, ensuring that only domestically active firms export imposes a limit on relative wages. The limit condition is discussed in the online appendix.

Note that we can formulate average revenues in terms of the cutoff productivities, $R_1(\tilde{\varphi}_1) = \left(\frac{\tilde{\varphi}_1}{\varphi_1^*}\right)^{\sigma-1} R(\varphi_1^*)$. By combing the profits from domestic and export sales with the conditional export probability in Eq. (6), the zero-cutoff-profit condition can be derived as

$$\bar{\pi}_1 = \left(\frac{\tilde{\varphi}_1}{\varphi_1^*}\right)^{\sigma-1} f w_1 - f w_1 + \left(\frac{\varphi_1^*}{\varphi_{1x}^*}\right)^k \left[\left(\frac{\tilde{\varphi}_{1x}}{\varphi_{1x}^*}\right)^{\sigma-1} f_x w_1 - f_x w_1 \right], \tag{7}$$

where the first two terms in the RHS are domestic profit whereas the third one is the profit from export market. Then, by combining Eqs. (4) and (7), we solve the domestic cutoff level of productivity in region 1 as

$$\varphi_1^* = \left[\frac{f(\sigma-1)}{f_e(k-\sigma+1)} \frac{1-\mathcal{H}^2\phi^2}{1-\mathcal{H}\phi(w_2/w_1)^{\frac{k\sigma}{\sigma-1}}} \right]^{1/k}, \tag{8}$$

where $\mathcal{H} \equiv (f_x/f)^{\frac{k-\sigma+1}{1-\sigma}} \in (0, 1)$.⁴ It shows that the region with higher wages features lower cutoff productivity because higher wages reduce expected profits and result in less entry.

Finally, we close the model by using the labor market clearing condition jointly with expenditure balance condition. Specifically, the labor market clearing condition of region 1 can be formulated as

$$L\lambda = n_1 \left(\frac{d_1(\tilde{\varphi}_1)}{\tilde{\varphi}_1} + f \right) + n_1 \left(\frac{\varphi_1^*}{\varphi_{1x}^*} \right)^k \left(\frac{\tau d_{1x}(\tilde{\varphi}_{1x})}{\tilde{\varphi}_{1x}} + f_x \right) + (\varphi_1^*)^k n_1 f_e, \tag{9}$$

in which n_1 is the number of varieties and the demands are functions of the price index given by

$$P_1 = \frac{\sigma}{\sigma-1} \left[n_1 \left(\frac{w_1}{\tilde{\varphi}_1} \right)^{1-\sigma} + n_2 \left(\frac{\varphi_2^*}{\varphi_{2x}^*} \right)^k \left(\frac{\tau w_2}{\tilde{\varphi}_{2x}} \right)^{1-\sigma} \right]^{\frac{1}{1-\sigma}}.$$

The LHS of Eq. (9) is the total labor supply in region 1, whereas the RHS represents the labor employed in domestic production, export and fixed entry inputs, respectively. On the other hand, the total expenditures in domestic and foreign market equate the total revenues. We thus have

$$E_1(1-\beta) = n_1 \left(\frac{\tilde{\varphi}_1}{\varphi_1^*} \right)^{\sigma-1} \sigma f w_1 + n_2 \left(\frac{\varphi_2^*}{\varphi_{2x}^*} \right)^k \left(\frac{\tilde{\varphi}_{2x}}{\varphi_{2x}^*} \right)^{\sigma-1} \sigma f_x w_2, \tag{10}$$

where the LHS is total expenditure on manufacturing goods while the RHS are revenues of domestic and foreign manufacturing firms, respectively. Note that, for

⁴ Assuming a Pareto distribution implies that average productivity results as a constant markup over the respective cutoff levels, that is $\tilde{\varphi}_1/\varphi_1^* = \tilde{\varphi}_{1x}/\varphi_{1x}^* = [k/(k-\sigma+1)]^{1/(\sigma-1)}$, which helps to simplify the mathematical expressions.

Eqs. (9), (10), mirror expressions exist for region 2, and we thus have four equations that endogenously determine the variables, w_1, w_2, n_1 and n_2 .

4 The spatial equilibrium

In the long-run, workers migrate across regions driven by utility differential. As in the literature, migration is governed by the migration equation:

$$\dot{\lambda} = (v_1 - v_2)(1 - \lambda)\lambda,$$

where v_r is the utility level of a representative worker in region r . It demonstrates that the long-run equilibrium holds when all workers agglomerate in one region ($\lambda = 1$ or 0) or $v_1 - v_2 = 0$ with $\lambda \in (0, 1)$. As in Krugman’s core-periphery model, endogenous variables enter in a non-linear fashion such that closed-form solutions for the full range of spatial distribution of workers are generally infeasible. Nevertheless, we can solve the model analytically for the symmetric equilibrium ($\lambda = 1/2$) and full agglomeration ($\lambda = 1$ or 0). To rule out the first nature difference, as in Helpman (1998), the housing stocks are assumed to be the same across regions.⁵ In region 1, per capita consumption of housing equals $S_1/L\lambda$ in which S_1 is the fixed housing stock. By using Eq. (1), the utility level of a representative worker in region 1 is derived as

$$v_1 = \left(\frac{S_1}{L\lambda}\right)^\beta \left[\frac{E_1(1 - \beta)}{L\lambda P_1}\right]^{1-\beta}. \tag{11}$$

4.1 Symmetry

At symmetric equilibrium, the population is equally distributed across regions ($\lambda = 1/2$). By using Eqs. (9)–(10) and the corresponding mirror equations, the wages and equilibrium firm numbers are uniquely solved as:

$$w_1 = w_2 \equiv w, \quad n_1 = n_2 = \frac{L(k - \sigma + 1)}{2fk\sigma(1 + \mathcal{H}\phi)}.$$

As in the literature, the symmetric equilibrium is stable⁶ if and only if $\frac{\partial(v_1/v_2)}{\partial\lambda}\Big|_{\lambda=1/2} < 0$. Using the total differentials of Eqs. (9)–(10) and the corresponding mirror equations, jointly with the equilibrium values of wages and firm numbers at $\lambda = 1/2$, we derive:

$$\frac{\partial(v_1/v_2)}{\partial\lambda}\Big|_{\lambda=1/2} = -\frac{4\mathcal{F}(\phi)}{(\sigma - 1)^2[(2 - \beta)\mathcal{H}^2\phi^2 + 2\left(\frac{2k\sigma - \sigma + 1}{\sigma - 1}\right)\mathcal{H}\phi + \beta]}, \tag{12}$$

where $\mathcal{F}(\phi)$ is defined as $\mathcal{F}(\phi) \equiv \mathcal{A}\mathcal{H}^2\phi^2 + \mathcal{B}\mathcal{H}\phi + \mathcal{C}$, with

⁵ We examine the case of asymmetric housing stocks in Sect. 4.4.

⁶ Readers who are interested in stability analysis of spatial models with dynamics could reference, for example, Neto and Claeysen (2015).

$$\begin{aligned} \mathcal{A} &\equiv (2 - \beta)(\sigma - 1)[(2 - \beta)\sigma - 1] > 0, \\ \mathcal{B} &\equiv 2[-\sigma(\sigma - 1)\beta^2 + 2k\sigma^2\beta + \sigma - 1 - 2k\sigma], \\ \mathcal{C} &\equiv \beta(\sigma - 1)(\beta\sigma - 1). \end{aligned}$$

Equation (12) shows that the symmetric equilibrium is stable if and only if $\mathcal{F}(\phi) > 0$. To simplify the expression, $\mathcal{F}(1)$ is rearranged as a quadratic function in term of β :

$$\mathcal{F}(1) = \mathcal{Q}_1\beta^2 + \mathcal{Q}_2\beta + \mathcal{Q}_3,$$

where

$$\begin{aligned} \mathcal{Q}_1 &\equiv (1 - \mathcal{H})^2(\sigma - 1)\sigma > 0, \\ \mathcal{Q}_2 &\equiv 4\mathcal{H}k\sigma^2 - (\sigma - 1)(4\mathcal{H}^2\sigma - \mathcal{H}^2 + 1), \\ \mathcal{Q}_3 &\equiv 2\mathcal{H}[(\sigma - 1)(2\sigma\mathcal{H} + 1 - \mathcal{H}) - 2k\sigma] < 0. \end{aligned}$$

Lemma 1 Define $\beta^\sharp \equiv \frac{-\mathcal{Q}_2 + \sqrt{\mathcal{Q}_2^2 - 4\mathcal{Q}_1\mathcal{Q}_3}}{2\mathcal{Q}_1}$, and we have $\beta^\sharp < 1/\sigma$. If $0 < \beta < \beta^\sharp$, we have $\mathcal{F}(1) < 0$; if $\beta > \beta^\sharp$, we have $\mathcal{F}(1) > 0$.

Proof $\mathcal{F}(1)$ is a quadratic function in terms of β . At $\beta = 0$, we have $\mathcal{F}(1) = \mathcal{Q}_3 < 0$; on the other hand, at $\beta = 1/\sigma$, we have $\mathcal{F}(1) = 2\mathcal{H}(\sigma - 1)^2(2\sigma\mathcal{H} + 1 - \mathcal{H})/\sigma > 0$. Moreover, $\mathcal{Q}_1 > 0$ implies that $\mathcal{F}(1)$ is a convex function in terms of β . Therefore, there exists a unique $\beta^\sharp \in (0, 1/\sigma)$ at which $\mathcal{F}(1) = 0$. If $\beta < \beta^\sharp$, we have $\mathcal{F}(1) < 0$; if $\beta > \beta^\sharp$, we have $\mathcal{F}(1) > 0$. \square

Based on the results above, we solve the threshold value of transport costs and obtain a proposition as follow.

Proposition 1 If $\beta > 1/\sigma$, the symmetric equilibrium is always stable; if $\beta < \beta^\sharp$, the symmetric equilibrium is always unstable; if $\beta^\sharp < \beta < 1/\sigma$, there exists a unique τ -break point given by

$$\tau_b \equiv \left(\frac{-\mathcal{B}\mathcal{H} + \sqrt{\mathcal{B}^2\mathcal{H}^2 - 4\mathcal{A}\mathcal{H}^2\mathcal{C}}}{2\mathcal{A}\mathcal{H}^2} \right)^{-1/k} > 1,$$

below which the symmetric equilibrium is stable.

Proof See ‘‘Appendix 1’’. \square

Contrary to the results of the standard core-periphery model, the symmetric equilibrium is stable when the transport costs are sufficiently low (i.e., $\tau < \tau^b$). The difference comes from the centrifugal forces. In the core-periphery model, the centrifugal force originates from the agricultural sector whose share in employment and expenditure has sharply decreased in most industrialized countries, as argued by Murata and Thisse (2005, p. 138). In contrast, the main centrifugal force here stems from the housing congestion, which fits well with modern urban-space economies.

For a given housing stock, agglomeration brings to lower per capita housing consumption. Thus, mobile workers disperse in order to alleviate the congestion when transport costs are sufficient low. In particular, the results show that, if the housing preference is extremely high (i.e., $\beta > 1/\sigma$),⁷ the centrifugal force is so strong that the symmetric equilibrium is always stable. On the other hand, if the housing preference is extremely low (i.e., $\beta < \beta^{\#}$), residents care little about housing congestion and, as a result, the symmetric equilibrium is always unstable.

The results are consistent with those of Helpman (1998), Tabuchi (1998), and Murata and Thisse (2005). However, by introducing firm heterogeneity *à la* Melitz (2003) into the model of Helpman (1998), our contributions are manifold. First, while the main findings in Helpman (1998) are derived through numerical simulations, we analytically give explicit solutions for the threshold values of housing preference β and transport costs τ_b . Second, by incorporating firm heterogeneity in productivity, our setup is, no doubt, more comprehensive and general. We analytically show that the main findings of Helpman (1998) are robust even if firm heterogeneity in productivity is considered. Third, Helpman (1998) is a highly cited paper and, in particular, its setup has also been widely employed in the recent literature on quantitative spatial economics (see, e.g., Redding 2016 and Redding and Rossi-Hansberg 2017). Our further exploration of this framework extends our knowledge about its theoretical foundation. Moreover, our setup enables us to explore how a change in firm heterogeneity (i.e., a fall of k) affects economic agglomeration in an urban-space economy, which is explored in more detail in the following subsections.

4.2 Full agglomeration

This section examines the equilibrium of full agglomeration ($\lambda = 1$).⁸ As in the literature, the full agglomeration is sustainable if and only if $\frac{v_1}{v_2} \Big|_{\lambda=1} > 1$. Plugging $\lambda = 1$ into Eq. (10) gives n_1 , and at λ close to 1, by Eq. (9), we solve

$$n_1 = \frac{L(k - \sigma + 1)}{fk\sigma}, \quad \left(\frac{w_1}{w_2}\right)^{\frac{k\sigma}{1-\sigma}} = \frac{k(1 - \lambda)[\beta(\sigma - 1) - 1]}{\mathcal{H}\phi\{\sigma - 1 + k(1 - \lambda)[\beta\lambda(\sigma - 1) - 1]\}} + \mathcal{H}\phi. \quad (13)$$

By the definition of v_r (Eq. (11)), Eqs. (5), (8), (13), we derive

⁷ Note that, if σ is high, consumers view different varieties as closer substitutes and care little about the available variety choice, which enlarges the range of stable symmetric equilibrium. As shown in “Appendix 2”, a higher σ decreases the range of stable full agglomeration.

⁸ The scenario of $\lambda = 0$ can be derived analogously.

$$\begin{aligned}
 \frac{v_1}{v_2} \Big|_{\lambda \rightarrow 1} &= \underbrace{\left[\frac{n_2 \left(\frac{w_2}{\varphi_2}\right)^{1-\sigma} + n_1 \left(\frac{\varphi_1^*}{\varphi_{1x}}\right)^k \left(\frac{\tau w_1}{\varphi_{1x}}\right)^{1-\sigma}}{n_1 \left(\frac{w_1}{\varphi_1}\right)^{1-\sigma} + n_2 \left(\frac{\varphi_2^*}{\varphi_{2x}}\right)^k \left(\frac{\tau w_2}{\varphi_{2x}}\right)^{1-\sigma}} \right]^{\frac{1-\beta}{1-\sigma}}}_{\text{price index effect}} \left[\frac{w_1 + \left(\frac{\beta}{1-\beta}\right)(\lambda w_1 + (1-\lambda)w_2)}{w_2 + \left(\frac{\beta}{1-\beta}\right)(\lambda w_1 + (1-\lambda)w_2)} \right]^{1-\beta} \\
 &\times \underbrace{\left(\frac{1-\lambda}{\lambda}\right)^\beta}_{\text{congestion effect}} \Big|_{\lambda \rightarrow 1} \\
 &= \Psi(\tau)(1-\lambda)^{\frac{\beta(k\sigma-\sigma+1)-(k-\sigma+1)}{k(\sigma-1)}} \Big|_{\lambda \rightarrow 1},
 \end{aligned}
 \tag{14}$$

where

$$\Psi(\tau) \equiv \left[\frac{(1-\mathcal{H}^2\tau^{-2k})(\sigma-1)}{k(\sigma\beta-\beta-1)} \right]^{\frac{(1-\beta)(k-\sigma+1)}{k(\sigma-1)}} \left[(\mathcal{H}\tau^{-k})^{\frac{2\sigma-1}{k\sigma}}(1-\beta) + \beta(\mathcal{H}\tau^{-k})^{\frac{1}{k}} \right]^{\beta-1}.$$

Equation (14) shows that, as workers fully agglomerate into one region, all varieties are free of transport costs and, as a result, the local price index becomes extremely low (relative to the periphery region), as captured by the *price index effect*. On the other hand, as λ approaches one, the per capita housing consumption goes extremely low, which discourages further agglomeration, as captured by the *congestion effect*. Evidently, there is a trade-off between these two effects and, in particular, this trade-off is reflected by the power exponents of β , σ , and k . Further analytical investigation gives us the following proposition.

Proposition 2 *If $\beta > \frac{k-\sigma+1}{k\sigma-\sigma+1}$, full agglomeration is always unstable; if $\beta < \frac{k-\sigma+1}{k\sigma-\sigma+1}$, full agglomeration is always stable; if $\beta = \frac{k-\sigma+1}{k\sigma-\sigma+1}$, a sustain point is defined by the equation $\Psi(\tau_s) = 1$, beyond which the full agglomeration is sustainable. Moreover, an increase in firm heterogeneity decreases the range in which full agglomeration is stable.*

Proof See “Appendix 2”. □

As shown by the above results, the stabilities of the full agglomeration depend on the trade-off between the *price index effect* and the *congestion effect*. These two effects in turn crucially depend on the power exponents of β , σ , and k . We analytically show that, if the housing preference is high (i.e., $\beta > \frac{k-\sigma+1}{k\sigma-\sigma+1}$), full agglomeration is always unstable. In other words, if the housing preference is particularly high, the *congestion effect* dominates, and the full agglomeration collapses if one worker leaves the core. On the other hand, if the housing preference is low (i.e., $\beta < \frac{k-\sigma+1}{k\sigma-\sigma+1}$), the *price index effect* dominates and, as a result, workers fully agglomerate to take advantage of the larger market.

Helpman (1998) finds that the full agglomeration is sustainable if the elasticity of substitution and the intensity of preferences for housing are small (i.e., $\beta\sigma < 1$) by numerical simulations. In contrast, this paper contributes to the literature by analytically giving the explicit solution for the threshold value of β below which the full agglomeration is sustainable. Moreover, in a more comprehensive setup with firm heterogeneity in productivity, we analytically disclose that, besides β and σ , the measure of firm heterogeneity (i.e., k) enters the expression of the threshold value of β as well. Our comparative static analysis shows that a fall in k decreases the range of β in which the full agglomeration is sustainable. In other words, an increase of firm heterogeneity works in favor of dispersion.

Intuitively, as in the literature (e.g., Ehrlich and Seidel 2013), a change in firm heterogeneity affects the centrifugal and centripetal forces through several channels. First, an increase in firm heterogeneity implies a larger proportion of high-productivity firms and a higher propensity to export. As a larger proportion of globally produced varieties becomes available and can be imported at lower prices, living in the larger market becomes less important. It attenuates the centripetal force originating from the *price index effect*. Second, a fall of k means that it becomes more likely to draw high productivity levels, which in turn implies lower prices and a higher propensity to export. With a larger proportion of high-productivity competitors, firms in the larger market are shielded less from their competitors. As a result, the profits earned from the local market fall, rendering the larger market less attractive than before, which also attenuates the centripetal force. Third, an increase in firm heterogeneity also implies fewer but more efficient firms. Each incumbent in the market owns a larger market share such that the immigration of one worker to the larger market reduces the market share of each incumbent relatively less. In this way, the centrifugal force originating from competition also decreases in magnitude. The former two outweigh the third one, and an increase in firm heterogeneity works in favor of dispersion.

In this regard, our results agree with the findings by Behrens et al. (2011), revealing that firm heterogeneity acts as a dispersion force, as the more productive firms crowd their less productive rivals out of the larger market. Our results are also consistent with those of Zhou (2018), showing that an increase of firm heterogeneity fosters dispersion in a linear city model with commuting costs and land rent.

To better understand the impact of trade liberalization or a change in firm heterogeneity on spatial equilibrium, we classify the results of related literature and summarize them into Table 1.

It is interesting to find that, in each different context, trade liberalization or an increase in firm heterogeneity affects spatial equilibrium in the same directions. Intuitively, we have explained how an increase in firm heterogeneity affects the centrifugal and centripetal forces through the aforementioned channels. It is noteworthy that a fall of transport costs works in a similar way. Specifically, as the transport costs fall, (i) varieties produced in the larger market can be imported at lower prices, which makes living in the larger market less important and attenuates the *price index effect*; (ii) firms in the larger market are shielded less from foreign competitors, rendering the *home-market effect* less attractive; (iii) firms sell more to the foreign market and the immigration of one worker to the larger market reduces

Table 1 The spatial impact of trade liberalization or a change in firm heterogeneity: a comparison of related literature

Effect \ Model	Helpman (1998) & This article	Murata and Thisse (2005) & Zhou (2018)	Krugman (1991a) & Ehrlich and Seidel (2013)
A fall of transport costs	Dispersion	Dispersion	Agglomeration
An increase in firm heterogeneity	Dispersion	Dispersion	Agglomeration

the market share of each incumbent relatively less (*competition effect*).⁹ Although the magnitudes of the impacts on each effect may be quite different, in each context, as shown by Table 1, an increase in firm heterogeneity affects the spatial equilibrium in the same directions as trade liberalization.

Does firm heterogeneity work in favor of agglomeration or dispersion? Existing literature provides ambiguous answers.¹⁰ We interpret that an increase in firm heterogeneity works in a very similar way as trade liberalization. In turn, the impacts of trade liberalization depend on the specific dispersion forces in each context. For instance, trade liberalization tends to elicit the equilibrium of full agglomeration in the model of Krugman (1991a) in which the dispersion forces originate from the expenditures of the immobile farmers. In contrast, in Helpman (1998) and Murata and Thisse (2005), trade liberalization works in favor of dispersion, because workers and firms tend to disperse in order to alleviate the urban costs when transport costs across regions are low.¹¹

To close the discussions in this subsection, we summarize the stabilities of the equilibria by arranging the threshold values of β . Plugging $\beta = \frac{k-\sigma+1}{k\sigma-\sigma+1}$ into $\mathcal{F}(1)$ gives

$$\mathcal{F}(1) = -\frac{(1-\mathcal{H})(\sigma-1)^2}{(k\sigma-\sigma+1)^2} [(k-\sigma+1)(\sigma-1) + \mathcal{H}(2k\sigma-\sigma+1)(2k\sigma-k-\sigma+1)] < 0,$$

where the inequality comes from $k > \sigma - 1$. By the properties of $\mathcal{F}(1)$, we have $\frac{k-\sigma+1}{k\sigma-\sigma+1} < \beta^{\#}$. The stabilities of the equilibria w.r.t. the values of β are summarized in the Table 2.

⁹ In our model with firm heterogeneity, trade liberalization also works through an exporter selection effect. Lower transport costs allow less productive firms to earn positive profits from foreign markets, which implies a higher propensity to export and raises the share of available varieties in each market. At the same time, intensified competition from high-productive firms drives the least productive firms out of the market and raises the output and market share of each incumbent. As also mentioned by Ehrlich and Seidel (2013, p. 543), “Firm heterogeneity works in a very similar way as the selection effects that come along with trade liberalization.”

¹⁰ Ehrlich and Seidel (2013) find that an increase in firm heterogeneity fosters agglomeration while Behrens et al. (2011) and Zhou (2018) provide the opposite results. Baldwin and Okubo (2006) also argue that firm heterogeneity can be thought of as a dispersion force in the sense that a smaller share of firms will have relocated from the small to the large region. Among others, Okubo et al. (2010, p. 231) argue that heterogeneity may act as an agglomeration force or as a dispersion force.

¹¹ Readers who are familiar with the literature on trade and urbanization may remember the fact that trade liberalization leads to dispersion across regions in Krugman and Elizondo (1996) in which centrifugal force mainly originates from urban costs, while Paluzie (2001) gives the opposite result with centrifugal force stemming from the immobile expenditure.

Table 2 The stabilities of equilibria w.r.t. the values of β

β	$(0, \frac{k-\sigma+1}{k\sigma-\sigma+1})$	$(\frac{k-\sigma+1}{k\sigma-\sigma+1}, \beta^d)$	$(\beta^d, 1/\sigma)$	$(1/\sigma, 1)$
Symmetry	Unstable	Unstable	Stable if $\tau < \tau_b$	Stable
Full agglomeration	Stable	Unstable	Unstable	Unstable

4.3 The set of equilibria: numerical examples

The forgoing analyses, however, do not provide a full characterization of spatial equilibria. Since the closed-form solutions for the full range of spatial distribution of workers are difficult, we appeal to numerical experiments in this section.

Figure 1 is depicted for parameters chosen as $\sigma = 3$, $f = 5$, $f_x = 35$ and $f_e = 1$ that are in line with the recent literature (Egger et al. 2013; Ehrlich and Seidel 2015).¹² The vertical axis represents the utility differential, v_1/v_2 , while the horizontal axis represents the population share. The rows show the results of the given β and decreasing transport costs τ while the columns show those of the given τ and increasing β . The main numerical results may be summarized as follows.

First of all, it illustrates that, for any given levels of transport costs, a higher β fosters dispersion. To be specific, in the first row, if $\beta = 0.1 < \frac{k-\sigma+1}{\sigma k - \sigma + 1} (\geq 0.14)$, the full agglomeration is always stable, thus confirming the prediction of Proposition 2. Also, $\beta = 0.1 < \beta^d (\geq 0.25)$, as predicted by Proposition 1, the dispersion is always unstable. On the other hand, if $\beta = 0.6 > 1/\sigma (\approx 0.33)$, the third row shows that the symmetric equilibrium is always stable, confirming the prediction of Proposition 1. Also, if $\beta = 0.6 > \frac{k-\sigma+1}{k\sigma-\sigma+1} (\leq 0.23)$, as shown in the third row, full agglomeration is always unstable, thus confirming Proposition 2.

Second, the second row shows that, for a given β , lower transport costs favor the equilibrium of dispersion. Intuitively, lower transport costs make it unnecessary to concentrate together in order to alleviate the housing congestion. It is particularly obvious for cases with higher values of β . It is noteworthy that the second row shows the existences of stable partial equilibria. Even though housing stocks are the same across regions, it is interesting to find that the two regions have different sizes of population. In this equilibrium, the residents in the larger region consume a better variety choice which is just sufficient to compensate them for the lower per capita housing consumption. As a result, residents in the larger and smaller region have the same levels of utility.

Last but not least, in all the cases shown by Fig. 1, a fall of k , thus a higher level of firm heterogeneity works in favor of dispersion, confirming the former analytical results. The numerical results confirm our forgoing theoretical predictions.

¹² The main results are insensitive to to the choice of these parameters, and the results for alternative parameters are provided from the author upon request.

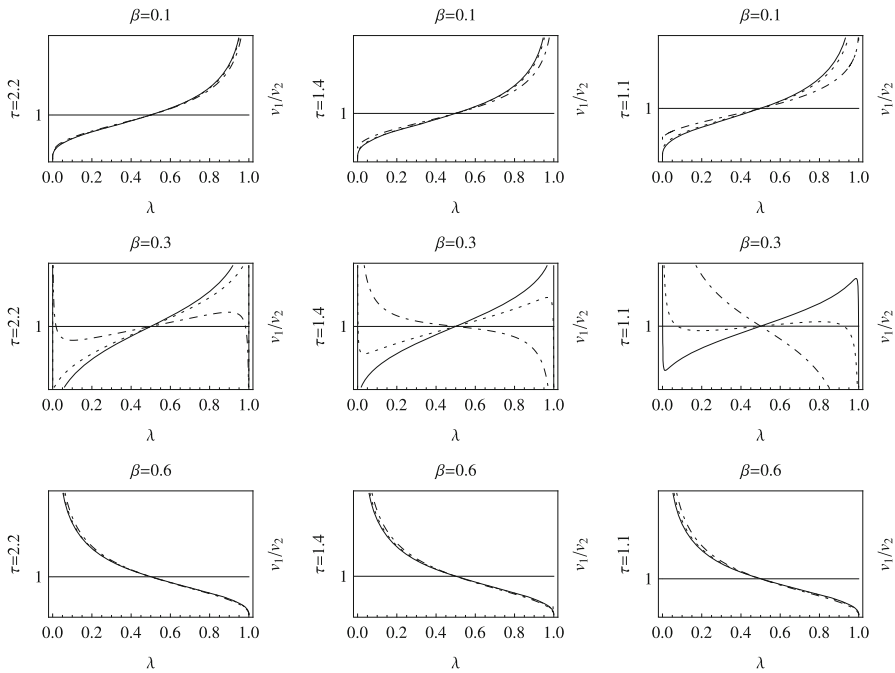


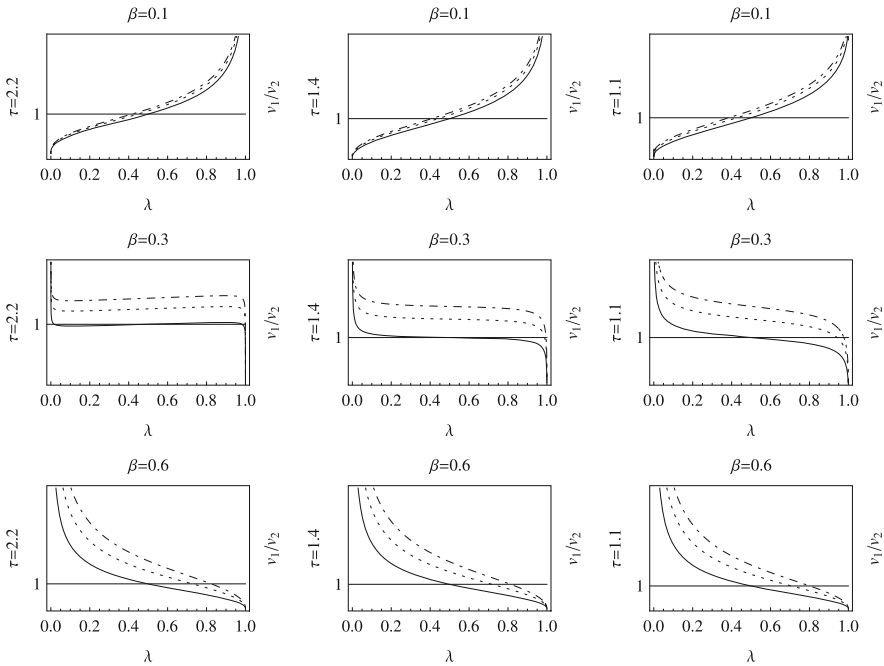
Fig. 1 Utility differentials and population distribution. Solid: $k = 5$, Dotted: $k = 4$, Dotdashed: $k = 3$

4.4 The case of asymmetric housing stocks

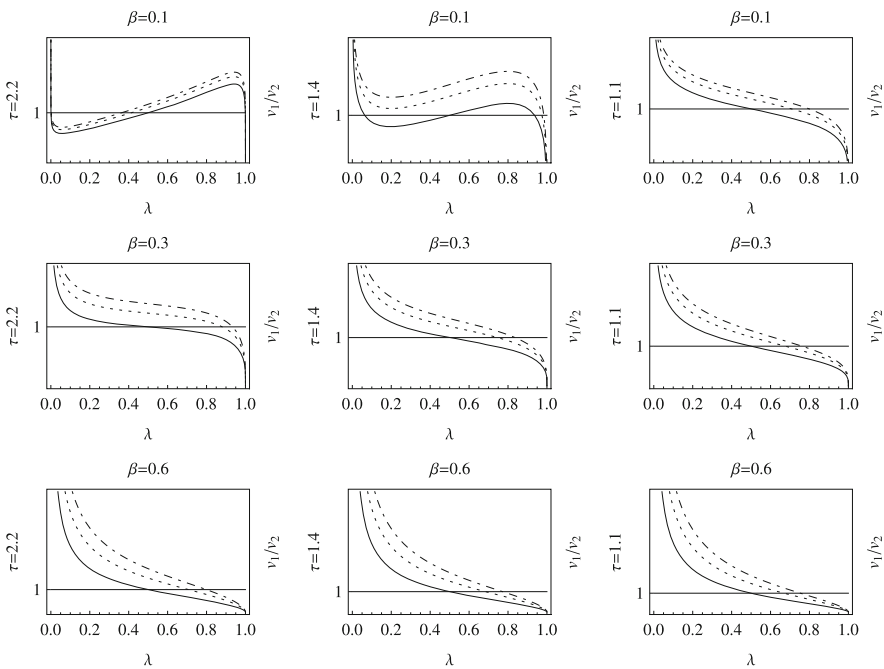
This subsection complements the results above by analyzing the case of asymmetric housing stocks. In the original work of Helpman (1998), the analysis of asymmetric settings is limited to extreme case of free trade. In contrast, we first extend his analysis to a more general context, and then compare the results in Helpman’s setup with those in our model with firm heterogeneity. Since the closed-form solutions are not available, we appeal to numerical experiments. The parameters are given as the same in Fig. 1 and k is chosen as $k = 3$. The solid, dotted and dotdashed curves in Fig. 2 depict the cases of $S_1/S_2 = 1$, $S_1/S_2 = 2$, and $S_1/S_2 = 3$, respectively.

Several comments are in order. First, as illustrated by Fig. 2, a higher housing preference (i.e., β) or a lower level of transport costs (i.e., τ) fosters dispersion. It indicates that the main predictions in symmetric settings remain largely robust when housing stocks are unevenly distributed across regions.

Second, Helpman (1998) predicts that when transport costs are very close to the level of free trade, the share of the population is *proportional* to each region’s housing stock. This prediction is confirmed by the third column in Fig. 2b, showing that if the housing stock in region 1 is twice that of region 2 (the dotted curve), the share of the population in region 1 is almost proportional ($\approx 0.66/0.33$). In contrast, in our model with firm heterogeneity, Fig. 2a illustrates that, even when the transport costs are negligible (i.e., the third column), the share of the population is *more than proportional* to the larger region’s housing stock. Intuitively, this



(a) Utility differentials with asymmetric housing stocks and firm heterogeneity.



(b) Utility differentials with asymmetric housing stocks in the setup of Helpman (1998)

Fig. 2 A comparison between Helpman (1998) and this article in asymmetric settings

difference arises from the firm heterogeneity of productivity. In the model of Helpman (1998), the firms are homogeneous in productivity. If the transport costs are negligible, each resident has access to all brands of varieties and, thus, her utility from the consumption of differentiated goods does not depend on the region in which she chooses to live. Her income is also the same in either of the regions in free trade. Therefore, the resident migrates to the region with lower housing costs until each region's population becomes proportional to its housing stock.

In contrast, in our setup, the firms are heterogeneous in productivity. Due to the existence of fixed export barriers, only a proportion of high productivity firms export, even when the transport costs are negligible. This gives the consumers in the larger better access to the differentiated goods, and this benefit from the home-market effect is just enough to compensate for the congestion resulting from a *more than proportional* share of the population in that region.

Third, in Fig. 2a, if the housing preference is low (i.e., $\beta = 0.1$), each resident cares little about the consumption of housing service, and the home-market effect always dominates. As a result, even if the housing stocks are unevenly distributed, the equilibrium of full agglomeration in either region is sustainable. In other words, *it is very possible that the region with more housing stock turns out to be the periphery region.*¹³ We note that there is an extensive body of literature¹⁴ documenting the phenomenon of “ghost cities” in urban development, especially in countries like China, where high-speed urbanization and real estate development are occurring. While the existing literature attributes the phenomenon to land-centred urbanization, excessive housing and infrastructure construction, and speculation on property demand, this paper provides alternative explanations based on a solid theoretical foundation.

5 Concluding remarks

By introducing Melitz-type firm heterogeneity into the model of Helpman (1998), we revisit his main findings by showing how the spatial equilibrium of an urban-space economy is determined by the trade-off between the centripetal force arising from the home-market effect and the centrifugal force originating from housing congestion. More importantly, since Helpman (1998) is a highly cited paper and its setup has been widely employed in recent literature on quantitative spatial economics, this study contributes to the literature by analytically providing explicit solutions for the threshold values of the housing preference (β) and the transport costs (i.e., τ_b , τ_s), which play crucial roles in determining the spatial equilibrium. In a more comprehensive setup with firm heterogeneity in productivity, this paper discloses analytically and intuitively how a change in firm heterogeneity affects the urban agglomeration. We compare our results with those of related studies and reveal that firm heterogeneity affects spatial equilibrium in a manner similar to trade

¹³ The region that has more people to begin with grows in size until it absorbs the entire population. That is, “history matters” (Krugman 1991b; Matsuyama 1991). In the setup of Helpman (1998), with asymmetric housing stocks, full agglomeration is also possible if β is extremely low and τ is very high.

¹⁴ See, e.g., Shepard (2015) and Chen et al. (2016), for reference.

liberalization. Moreover, in the case of uneven housing stocks, we obtain new results and explain intuitively how and why our results differ from those in the model of Helpman (1998). In particular, our results provide alternative theoretical explanations for phenomena such as “ghost cities” in countries undertaking high speed urbanization and real estate development.

Acknowledgements I am grateful to the editor and two anonymous referees for helpful comments and suggestions. The usual caveat applies. Financial support from the National Science Foundation of China (Grant Nos. 71663023, 71950001, 71773042) is gratefully acknowledged.

Appendices

Appendix 1: Proof of Proposition 1

If $\beta > 1/\sigma$, we have $\mathcal{C} > 0$. Multiplying the two roots which satisfy $\mathcal{F}(\phi) = 0$, the product is $\mathcal{C}/\mathcal{A} > 0$, which implies that the two roots have the same sign. Also, it is easy to check $\mathcal{B} > 0$ when $\beta > 1/\sigma$, which implies that the slop of $\mathcal{F}(\phi)$ at $\phi = 0$ is positive. Since $\mathcal{F}(\phi)$ is a convex function, we know that the two roots are both negative and $\mathcal{F}(\phi) > 0$ for $\phi \in (0, 1)$. By Eq. (12), it implies that the symmetric equilibrium is always stable.

If $\beta < \beta^\sharp < 1/\sigma$, we have $\mathcal{F}(0) = \mathcal{C} < 0$. Also, as shown by Lemma 1, if $\beta < \beta^\sharp$, we have $\mathcal{F}(1) < 0$. Note that $\mathcal{A} > 0$ and $\mathcal{F}(\phi)$ is a convex function. Due to the continuities, $\mathcal{F}(0) < 0$ and $\mathcal{F}(1) < 0$ imply $\mathcal{F}(\phi) < 0$ for $\phi \in (0, 1)$. By Eq. (12), it implies that the symmetric equilibrium is always unstable.

If $\beta^\sharp < \beta < 1/\sigma$, we have $\mathcal{F}(0) = \mathcal{C} < 0$ and $\mathcal{F}(1) > 0$ by Lemma 1. Because $\mathcal{F}(\phi)$ is a convex function, there exists a unique ϕ_b at which $\mathcal{F}(\phi_b) = 0$ and we have $\mathcal{F}(\phi) > 0$ when $\phi > \phi_b$. By the definition of ϕ_b , we solve the τ_b . □

Appendix 2: Proof of Proposition 2

If λ is very close to 1, $(1 - \lambda)$ is very close to zero. As shown in Eq. (14), if $\beta > \frac{k-\sigma+1}{k\sigma-\sigma+1}$, the power of $(1 - \lambda)$ is positive and, therefore, v_1/v_2 is very close to zero. As a result, the full agglomeration is unstable. On the other hand, if $\beta < \frac{k-\sigma+1}{k\sigma-\sigma+1}$, the power of $(1 - \lambda)$ is negative and v_1/v_2 is close to infinity, as a result, the full agglomeration is stable. If $\beta = \frac{k-\sigma+1}{k\sigma-\sigma+1}$, the term of $(1 - \lambda)^{\frac{\beta(k\sigma-\sigma+1)-(k-\sigma+1)}{k(\sigma-1)}}$ equates 1. The full agglomeration is sustainable if and only if $\Psi(\tau) > 1$. Also, it is easy to check that the $\Psi(\tau)$ is an increasing function in terms of τ and, therefore, the sustain point τ_s is uniquely defined by $\Psi(\tau) = 1$, beyond which the full agglomeration is sustainable. Moreover, we have

$$\partial \left(\frac{k - \sigma + 1}{k\sigma - \sigma + 1} \right) / \partial k = \frac{(\sigma - 1)^2}{(k\sigma - \sigma + 1)^2} > 0, \partial \left(\frac{k - \sigma + 1}{k\sigma - \sigma + 1} \right) / \partial \sigma = - \frac{k^2}{(k\sigma - \sigma + 1)^2} < 0,$$

implying that a smaller k enlarges the range of parameters in which the full

agglomeration is unstable while a smaller σ fosters the equilibrium of full agglomeration. □

References

- Alonso W (1964) Location and land use. Harvard University Press, Cambridge
- Baldwin RE, Okubo T (2006) Heterogeneous firms, agglomeration and economic geography: spatial selection and sorting. *J Econ Geogr* 6:323–346
- Behrens K, Mion G, Ottaviano GIP (2011) Economic integration and industry reallocations. In: Jovanovic MN (ed) International handbook on the economics of integration, vol 2. Edward Elgar, Cheltenham
- Chen M, Liu W, Lu D (2016) Challenges and the way forward in China's new-type urbanization. *Land Use Policy* 55:334–339
- Egger H, Egger P, Kreickemeier U (2013) Trade, wages, and profits. *Eur Econ Rev* 64:332–350
- Ehrlich M, Seidel T (2013) More similar firms-more similar regions? On the role of firm heterogeneity for agglomeration. *Reg Sci Urban Econ* 43:539–548
- Ehrlich M, Seidel T (2015) Regional implications of financial market development: industry location and income inequality. *Eur Econ Rev* 73:85–102
- Helpman E (1998) The size of regions. In: D P, Sadka E, Zilcha I (eds) Topics in public economics. Cambridge University Press, Cambridge
- Helpman E, Melitz M, Yeaple SR (2004) Export versus FDI with heterogeneous firms. *Am Econ Rev* 94:300–316
- Krugman P (1991a) Increasing returns and economic geography. *J Polit Econ* 99:483–499
- Krugman P (1991b) History versus expectations. *Q J Econ* 106:651–667
- Krugman P, Elizondo RL (1996) Trade policy and the Third World metropolis. *J Dev Econ* 49:137–150
- Martin P, Rogers CA (1995) Industrial location and public infrastructure. *J Int Econ* 39:335–351
- Matsuyama K (1991) Increasing returns, industrialization, and indeterminacy of equilibrium. *Q J Econ* 106:617–650
- Melitz M (2003) The impact of trade on intra-industry reallocations and aggregate industry productivity. *Econometrica* 71:1695–1725
- Murata Y, Thisse J (2005) A simple model of economic geography *à la* Helpman–Tabuchi. *J Urban Econ* 58:137–155
- Neto JPY, Claeysen JCR (2015) Capital-induced labor migration in a spatial Solow model. *J Econ* 115:25–47
- Nocke V (2006) A gap for me: entrepreneurs and entry. *J Eur Econ Assoc* 4:929–955
- Okubo T (2009) Trade liberalization and agglomeration with firm heterogeneity: forward and backward linkages. *Reg Sci Urban Econ* 39:530–541
- Okubo T, Picard PM, Thisse JF (2010) The spatial selection of heterogeneous firms. *J Int Econ* 82:230–237
- Ottaviano GIP (2011) New new economic geography: firm heterogeneity and agglomeration economies. *J Econ Geogr* 11:231–240
- Paluzie E (2001) Trade policy and regional inequalities. *Pap Reg Sci* 80:67–85
- Pflüger M, Tabuchi T (2010) The size of regions with land use for production. *Reg Sci Urban Econ* 40:481–489
- Proost S, Thisse JF (2019) What can be learned from spatial economics. *J Econ Lit* 57(3):575–643
- Redding SJ (2016) Goods trade, factor mobility and welfare. *J Int Econ* 101:148–167
- Redding SJ, Rossi-Hansberg E (2017) Quantitative spatial economics. *Ann Rev Econ* 9:21–58
- Saito H (2015) Firm heterogeneity, multiplant choice, and agglomeration. *J Reg Sci* 55(4):540–559
- Samuelson P (1954) The transfer problem and transport costs, II: analysis of trade impediments. *Econ J* 64:264–289
- Shepard W (2015) Ghost cities of China: the story of cities without people in the world's most populated country. Zed Books, London

- Tabuchi T (1998) Urban agglomeration and dispersion: a synthesis of Alonso and Krugman. *J Urban Econ* 44:333–351
- Zhou Y (2018) Heterogeneous firms, urban costs and agglomeration. *Int J Econ Theory*. <https://doi.org/10.1111/ijet.12194>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.