Humanities & Social Sciences Communications

ARTICLE

https://doi.org/10.1057/s41599-024-03627-9 OPEN

A Lucas island model to analyse labour movement choice between cities based on personal characteristics

Tiange Qi[1], Yuning Gao[2] & Yongjian Huang[3]✉

The labour movement has been a key factor for cities' development and caused regional inequality between cities. Although empirical studies have been conducted to investigate it, little theoretical evidence has been provided to find out the underlying mechanism. This paper describes and derives a Lucas-Prescott style island model to study the location choices of the heterogeneous agents by utilising endogenous technology growth, which in turn influences personal human capital growth. It leads to the U-shape curve of the inequality of wage income with the technology of these islands but not in terms of total income. In the extended two-goods model, the magnitude of the implications is increased by the impact of non-tradable goods price. Together with empirical research using the US census data, this paper finds that skilled labours with less endowed wealth tend to live in large cities for its high salary. On the other hand, those less-skilled but with more endowed wealth tend to live in cities with better environment, which drives up the price level of non-tradable goods in these cities. This explains the population concentration in the super cities and the high housing price-wage ratio in some beautiful cities, which provides theoretical basement for further empirical studies about labour movement in other cities

[1] Department of Economics, University of Cambridge, Cambridge, UK. [2] School of Public Policy and Management, Tsinghua University, Beijing, China. [3] School of Finance, Central University of Finance and Economics, Beijing, China. ✉email: 2022110050@email.cufe.edu.cn

Introduction and literature review

Since the advent of the Industrial Revolution, urbanisation has become a common trend across the world and has accelerated a great deal over recent years. According to the World Bank, database[1], the rate of worldwide urbanisation has risen from 34% to 37% from 1960 to 2022. This is underpinned by the fact that rates in developing countries such as Mexico have gone from 51% to 81% and in Brazil from 46% to 88% during that period. However, even developed countries such as US and UK have witnessed a rapid increase of the urbanisation rate, going from 79% to 83% and 79% to 84%, respectively within the last 20 years.

The concentration of population is a significant phenomenon for consideration during urbanisation. In specific countries, the population share of largest city to total urban population[2] is relatively high, such as in the UK (17%), Canada (20%), Mexico (21%), France (20%), Austria (23%), Japan (32%) and Chile (40%) when compared to Germany (5%) and US (7%). However, this percentage has been decreasing since 2000 in many major economies except for Austria and Canada, which have relatively lenient immigration policies. This population concentration brings about inequality, which is seldom addressed. Nord (1980) finds that with the increasing size of the city population, income equality produces a U-shaped curve, arguing that as the population immigrates to city, the occupational and wage structure of local labour market will change. Small cites lack infrastructure and sufficient size to support the types of business to distribute income. High-income business is monopolized at certain local enterprisers and immigrants have to enter the business with low income. As the city size increases, the improved infrastructure and demand provides capacity to different types of business and they provide higher salary to attract skilled immigrants, which would decrease inequality. However, Madden (2000) and Glaeser et al. (2009) point out inequality always increases with population. They find that in 2006, the elasticity of Metropolitan Statistical Area[3]'s Gini coefficient to city population is 0.036 and highly significant in US cities by US Census data. Additionally, they find that the mean derives from the top 5% mean income in a Metropolitan Statistical Area in relation to the overall mean income within that Metropolitan Statistical Area is highly significant. What's more, for high-productivity people, income is disproportionally positively correlated with the size of the city. They think that as population increases within the city, the market expands and the business, which requires high-skilled workers, will increase. This increases the demand for high-skilled workers and their wage return to human capital. At the same time, immigrants from Latin America, which are mainly low-skilled workers, worsens this wage inequality within the city. Wheeler (2001), Baum-Snow and Pavan (2009a) prove the robustness of this positive effect. The additional premium gained by a greater number of productive people arises in every quintile of the income distribution, as well as the development of the city has a direct effect on inequality even if the effect of socio-economic composition is ignored. Behrens and Robert-Nicoud (2014) build a model, which links the size of the city to inequality and productivity. In addition to the view that the increase of city size will increase the skill premium, they also argue that urbanization will improve competitiveness and openness to business, which decreases the cost of logistics and transportation and therefore increases the price competition locally (as the price of shipping goods to customers or their nearby shops is low, the local goods' advantages decrease). This will benefit the entrepreneurs which produce exported goods. And as only most productive companies export their goods, these companies will become richer, which increases inequality and productivity. Higher productivity will also increase the city size by attracting migrates, which is a cyclical effect. Castillo et al. (2020) notice that labour mobility will bring new knowledge to the companies, which benefits their productivity by knowledge diffusion, regardless of labour's level of skills.

The core to addressing these facts is understanding the drive behind different kinds of labour movements in different cities. Lucas and Prescott (1978) argue that decisions concerning labour movement should balance in terms of the trade-off for the benefit of moving to a higher-wage company (as island in the model) and the loss of wages during the period of 'searching for that company'. Therefore, they conclude that the technology heterogeneity between each island is the main drive behind unemployment as people have incentives to move to companies with higher technology to earn a higher wage, whereas those who are still searching for a better job contribute to unemployment. Then Coen-Pirani (2010) modifies the island model to include the use of land in the production function and land rent in income, applying it to analyse flow of labour between cities.

The theoretical rationale of this paper utilises the Lucas-Prescott model to address aspects of urbanisation, labour flow, and inequality. Rather than using a homogenised sense of labour in their model, this model introduces skill-heterogeneity between labour based on the model of Behrens and Robert-Nicoud (2014). Also, the model for this paper replaces land with capital as a complementing factor with labour to produce consumption goods. Additionally, both workers and capitalists are defined as agents. This is because land can be included in capital and it is also not the main factor for production in industrialised countries. In the UK, 57.5% of lands are used for the purpose of agriculture in 2018 (Ministry of Housing, Communities & Local Government, 2020), while the percentage of agriculture to UK GDP is only 0.61% (ONS, 2021). Introducing capital will add more flexibility and be more realistic. To address the facts of labour flow from large cities to cities that have beautiful natural landscapes, this paper also includes the utility of the quality of the environment, as it affects the movement of labour. Despite the wage mentioned above, another factor affecting job choice is environment (Feld et al., 2022). Chen et al. (2022) also find that the air pollution is responsible for large labour mobility in China: 10% air pollution will result in 2.8% population outflow in China in average. We therefore add the environmental factor in the labour's utility function.

Our paper makes contributions in several aspects. Firstly, is contributes to the theoretical literatures about labour's migration choice (Harris, 1970; Borjas, 1987; Zenou, 2008; Coen-Pirani, 2010; Chin and Cortes, 2015; Zhao, 2020; Aksoy and Poutvaara, 2021). Previous models address heterogeneous wage levels with homogeneous labours, such as McCall (1970) Research model where labours have a minimal wage to make them indifferent between accepting the job offer or waiting for a better job, which can also be used to explain unemployment rate. The Hopenhayn (1992) model addresses the entrance and exit mechanism of business in the labour market, of which the driving force is stochastic productivity of companies. The Aiyagari (1994) model analyses the matching process of labours with heterogeneous skill levels to homogeneous companies. The Krusell and Smith (1998) model combines the above two models, which incorporates a model with both heterogeneous labours and companies in terms of productivity. It finds the distribution of heterogeneous companies and their employees. However, these models do not consider the wealth as a factor in the matching process. We argue that the labour choice does not only depend on its human capital but also its asset, which complement their work, as the current research does not notice the negative effect of asset on the intention of labour to move for higher wage. Different from the mostly common search models, we introduce above where labour only care about consumption, our model is the first one to introduce environment to affect the labour choice. These two changes can explain the choices of different types of labours, not only in terms of the heterogeneous productivity but also the resources of the income to different cities.

Secondly, this paper constructs the distribution of labour in terms of workers' talent and their endowed capital in cities with degrees of different technology and environment, which provides a theoretical support to the empirical work of cities' wage differences, especially those of developing countries (Gong and Van Soest, 2002; Bargain and Kwenda, 2011; Buch et al., 2014; Cao et al., 2015). This is because the technology inequality between cities is server in developing countries and their industry structure means that technology development usually results in pollution, which will lead to higher rate of labour mobility.

Thirdly, this paper explains the price heterogeneity between cities due to the mobility of different labours. Van Nieuwerburgh and Weill (2010) propose an island model to analyse the effect of labour mobility on the housing price. They argue that the productivity is the determinant of labour mobility and wage level, which affect cities' housing price, which is consistent with our research. However, the positive correlation between wage and housing price do not apply in some cities, which will be discussed later. Our model with the environmental factor, asset and connection between the tradable goods and non-tradable goods solve this problem. We conclude that even though wage difference due to the technology inequality is positively connected with the price level of non-tradable goods, labour's mobility choice due to non-wage attributes plays an important role in the price level of non-tradable goods.

The remainder of this paper is organized as follows. Section "The baseline model" discusses the baseline model and establishes the equilibrium condition. And Section "Extended two-goods model" presents a more complicated version of the model, which allows for the inclusion of tradable and non-tradable goods to address the price of heterogeneity between different cities. While Section "Empirical analysis" carries out the simple empirical analysis, to support our model. Last, Section "Conclusion" concludes our paper.

The baseline model

Lucas-Prescott model. Within the Lucas-Prescott economy, people's income comes from their wages and the rent of capital. At the beginning of the period, each agent is on an island. The shocks are then publicly realised, and people respond to them by deciding to stay or leave the island. If people decide to leave the island, they do not work during the period (but can still receive the interest of the capital) and will go to a new island at the beginning of the next period. If they choose to stay, they can receive both wage income and the rent of capital. The island economy system is a closed economy without any international trade and capital inflow and outflow. Empirically, the system can be seen as a large economy or each island can be seen as a political identity. The monetary policy is neutral and there is no rigidity i.e. it only considers a real term.

A recursive equilibrium with rational and infinite life agents is used in this paper and is a set of functions:

$$\begin{gathered} v(\omega, z, h, \mu, q, k, \vartheta, K), e(\omega, z, h, \mu, q, k, \vartheta, K), \\ w(\omega, z, \mu, q, K), \omega'(\omega, z, \mu, q, K) \\ z'(z, \omega, q, K), h'(z, h), z'(z, \omega, q, K), r(\mu) \end{gathered},$$

where z is the technology in the island; q is the quality of living in each island or the function of the environment; h is the human capital of each individual; k is the capital owned by each individual and $\omega(h, k, \vartheta|z, q, K)$ is the number of people with human capital h, capital k and the speed of learning ϑ, which is the natural talent of each agent on the island with technology z, capital K and quality q.

$\mu(z, q, \omega, K)$is the density of the island with technology z, quality q and number of people ω and total capital K. $w(\omega, z, \mu, q, K)$ is the wage at the island and $r(\mu)$ is the rate of return to the capital. People can install their capital in each island without any barrier, which suggests that the return of capital for each island is the same. τ_w and τ_k are the tax rates of wage and capital return.

Each agent cares about both consumption and the quality with utility function. $u(c_t, q_t) = \ln(c_t) + \ln(q_t)$, wher c_t is the consumption at time t. People do not have disutility on labour and each one is endowed with 1 unit of labour in each period.

It satisfies:

$v(\omega, z, h, \mu, q, k, \vartheta, K) = \max\{v_{stay}(w, z, h, \mu, q, k, \vartheta, K), v_{leave}(k, h, \vartheta, \mu)\}$, which is the utility of the agent with h, ϑ, k at the island with ω z μ K q. He or she chooses the maximum between the utility of staying at the island or leaving the island at the start of each period.

Where

$$\begin{aligned} v_{stay}(\omega, z, h, \mu, q, k, \vartheta, K) = \max_{k'}\{ & u\big(h(1-\tau_w)w(\omega, z, \mu, q, k) + r(\mu)(1-\tau_k)k \\ & + (1-\delta_k)k - k', q\big) + \beta E_t v\big(\omega'(\omega, z, \mu, q, K, \vartheta), \\ & \times z'(z, \omega, q, K), h'(z, h), \mu'(\mu), k', q'(q, \mu)\big)\} \end{aligned}$$

is the utility of staying at the same island at the period. The agent chooses the investment to decide the capital next period k' to maximise the utility of consumption from return of the capital with depreciation rate $\delta_k \in (0,1)$ plus the wage as well as the utility of environment and the expected discounted utility at the next period with discount rate $\beta \in (0,1)$. The notation a' suggests the variable a at next period.

$$\begin{aligned} \nu_{leave}(k, h, \vartheta, \mu) = \max_{k'}\{ & u(r(\mu)(1-\tau_k)k + (1-\delta_k)k - k', \underline{q}). \\ & + \beta \max_{\tilde{\omega}, \tilde{q}, \tilde{z}, \tilde{K}} \{E_t \nu(\omega'(\tilde{\omega}, \tilde{q}, \tilde{z}, \tilde{K}, \mu), z'(\tilde{\omega}, \tilde{q}, \tilde{z}, \tilde{K}), \\ & \times h(1-\delta_h), \mu'(\mu), k', q'(\tilde{q}, \mu), \vartheta)\}\} \end{aligned}$$

is the utility of leaving the current island.

The agent is now on the way to search for new island. He or she chooses k' to get return from investment while the agent cannot receive the wage at the period and he or she enjoys the environment "on the way" $\underline{q}$, which is lower than the environment of all islands. The agent also decreases the human capital with rate $\delta_h \in (0,1)$.

$$\max_{\widetilde{\omega, q, z, K}} \{E_t \nu(\omega'(\widetilde{\omega}, \widetilde{q}, \widetilde{z}, \widetilde{K}, \mu), z'(\widetilde{\omega}, \widetilde{q}, \widetilde{z}, \widetilde{K}), h(1-\delta_h), \mu'(\mu), k', q'(\widetilde{q}, \mu), \vartheta)\}$$

suggests agent searches for an island to maximise the expected utility and arrives at that island at the start of next period.

z and h follow the growth path $\dot{z} = \theta(\int h * e(h, k, \vartheta|z, q, \omega, K) dh \times dk)^{\lambda} z^{\phi} z_n^{\gamma}$, which is the increase of technology and $\dot{h} = \vartheta h^{\eta} z^{\kappa}$ is the increase of the human capital.

Where $e(h, k, \vartheta|z, q, \omega, K)$ is the number of employed people with capital k, human capital h and speed of learning ϑ given the island with technology, quality, the number of people and total capital z, q, ω, K.

The structure of growth suggested by Romer (1986) will be used, which assumes that technology increase is the function of current technology, researchers in each island and 'natural technology' z_n, growing constantly at g. It is also assumed that the proportion of human capital used for the research is constant. This is the reason for introducing θ.

It will be applied to human capital growth. Human capital growth is positively correlated to current human capital and the technology of the city. The inclusion of technology is because the human capital need grow faster when the worker works in a company aiming with higher productivity (Abel et al., 2012). The elasticity of researchers to technology growth λ, the elasticity of current technology to technology growth ϕ, the elasticity of nature technology to technology growth γ, the elasticity current human capital to human capital growth κ and the elasticity of technology to human capital growth η are constant and within the range of (0,1).

There is a production sector in each island, which hires human capital and capital with production function *F(zH, K)*, which is homogeneous with degree 1, to solve:

Given $w(\omega, z, \mu, q)$ and $r(\mu)$

$$\max_{K,H}\left\{F(zH,K) - w\left(\omega, z, h, \mu, q, k\right)H - r\left(\mu\right)K\right\}$$

For consistency, each island i with z,q has:

$$H_i = \int h * e(k, h, \vartheta | z, q, \omega, K) dk dh d\vartheta$$

Equilibrium. Taking First order condition wrt K and H for each island i:

$$zF_1(zH, K) = w\left(\omega, z, \mu, q\right)$$

$$F_2(zH, K) = r(\mu)$$

The first order condition for capital is used to determine the allocation of capital in each island as $\forall$i F_2 (zH,K) and is equal for each period.

We can now find the market clearing condition:

For each island with technology and quality z and q, the agents with capital and human capital k and h can choose to

I. If $\nu_{stay}(\omega, z, h, \mu, q, k, \vartheta, K) \le \upsilon_{leave}(k, h, \vartheta, \mu)$, people with k and h leave the island and we have ν_{stay} $(\omega, z, h, \mu, q, k, \vartheta, K)$ $= \upsilon_{leave}(k, h, \vartheta, \mu)$ in equilibrium, people leave the island. Therefore, we have

$$\omega'\left(k', h', \vartheta, |, z'\left(\omega, q, z\right), q'\left(\mu\right), K'\right) = e(k, h, \vartheta | z, q, \omega)$$

II. If $\nu_{stay}(\omega, z, h, \mu, q, k, \vartheta, K) > \upsilon_{leave}(k, h, \vartheta, \mu)$ and

$$E_t\nu(\omega'(\omega, z, \mu, q, K, \vartheta), z'(z, \omega, q, K), h'(z, h), \mu'(\mu), k', q'(q, \mu)) < \max_{\tilde{\omega}, \tilde{q}, \tilde{z}, \tilde{K}}\{E_t\nu(\omega'(\tilde{\omega}, \tilde{q}, \tilde{z}, \tilde{K}, \mu), z'(\tilde{\omega}, \tilde{q}, \tilde{z}, \tilde{K}), h(1-\delta_h), \mu'(\mu), k', q'(\tilde{q}, \mu), \vartheta)\}\}$$

None of these people come and leave the island. Therefore, we have $\omega'(k', h', \vartheta | z'(\omega, q, z), q'(\mu), K') = \omega(k, h, \vartheta | z, q, K)$

III. If people come to the island and we can pin down ω' by the condition

$$E_t\nu(\omega'(\omega, z, \mu, q, K, \vartheta), z'(z, \omega, q, K), h'(z, h), \mu'(\mu), k', q'(q, \mu)) = \max_{\tilde{\omega}, \tilde{q}, \tilde{z}, \tilde{K}}\{E_t\nu(\omega'(\tilde{\omega}, \tilde{q}, \tilde{z}, \tilde{K}, \mu), z'(\tilde{\omega}, \tilde{q}, \tilde{z}, \tilde{K}), h(1-\delta_h), \mu'(\mu), k', q'(\tilde{q}, \mu), \vartheta)\}$$

The model also obtains:

$$\mu'\left(\mu\right)\left(\Delta z \times \Delta q \times \Delta\omega\right) = \frac{\int \mu(d\omega, dq, dz)}{z'(z, \omega) \in \Delta z, q' \in \Delta q, \omega'(\omega, z, \mu, q) \in \Delta\omega}$$

which is the density of the island's next period with technology, quality and distribution of capital and human capital Δz, Δq, $\Delta\omega$

Steady state. In the traditional Lucas island model, there is a stochastic item in the technology and there is still labour movement in the steady state. However, the distribution in terms of the islands' technology, production and labour is constant. However, as this model includes more complexity of technology and human capital growth, the existence of stochastic items will never reach a steady state in the model. Recall the growth of z depends on the z last period. If there is a stochastic item in the technology, the growth rate will never become constant. Therefore, for simplicity, our model does not include uncertainty and the steady state can simply be known as the population does change in each island, the technology and human capital grow as a constant rate and quality of environment is also constant. In the real world, it is apparently that it is impossible to reach such a steady state between cities because there are always labour with different levels of talents and wealth entering the market and retiring from the market. However, this model reveals a regulation of labour movement. Empirically, with the personal demography data, regressing their settlement choice with their demography, which is the topic of our future research also with deep empirical analysis, can prove this.

The steps to solving the steady state can be found in Appendix 1.

The condition for steady state can be solved now:

Condition 1 – For agents with initial value of k_0, h_0, ϑ at the time of steady state, they will only be on the island with:

IV. z_0, $q_0 \in \{z_0, q_0 | \nu_{stay}$ $(z_0, q_0, k_0, h_0, \vartheta) \ge \nu_{leave}$ $(k_0, h_0, \vartheta)\}$

Condition 2 – The sum of capital in each island should equal the sum of capital of each agent:

$$\iint k_0\omega(dk_0, dh_0, d\vartheta | z_0, q_0, K_0)\mu\left(dz_0, dq_0, dK_0\right) = \int K_0(z_0)\mu(dz_0) = \int k_0\xi(dk_0, d\vartheta),$$

where ξ is the pre-defined unconditional distribution of k_0 and ϑ.

Condition 3 – With the same logic, for each island i, the human capital employed by the company should be equal to the sum of human capital in that island:

$$\int h_0(z_0, \vartheta)\omega(k_0, h_0(z_0, \vartheta) | z_0, q_0))d\vartheta = H_0(z_0)$$

Condition 4 $\frac{K_t}{z_t H_t}$ is constant which can be solved by (3)

These four conditions suggest that given ξ and $\mu(z_0, q_0)$, there may be many steady state equilibriums if these conditions are satisfied. This is because condition 1 implies that for each pair of (k_0, ϑ), we can always find the set of (z_0, q_0) of which island the agent can stay in. There can be many allocations of each agent to the corresponding islands if the remaining conditions are satisfied.

Analysis. Switching from island A to island B increases the utility of quality $\Delta u_q = \ln(\underline{q}) + \frac{\beta \ln(q_B)}{1-\beta} - \frac{\ln(q_A)}{1-\beta}$, which is constant. In terms of the utility of consumption, switching from one island to another with higher technology will decrease human capital and the level of equilibrium $h_0(z_B, \vartheta) > h_0(z_A, \vartheta)$ at the time when the agent arrives at island B. Therefore, the agent needs time to reach the new equilibrium. For the human capital when the agent arrives at the island B, $\dot{h} = \vartheta h^\eta {z_B}^\kappa$, we use the first order Taylor approximation:

$$\dot{h} = \vartheta h^\eta {z_B}^\kappa = \dot{h}^* + \vartheta\eta {h^*}^{\eta-1}{z_B}^\kappa(h - h^*)$$

Where h^* is steady state level of h (z_B, ϑ) at the current time.

Multiply $\frac{1}{g_h h}$ and substitute $\vartheta {h^*}^{\eta-1} {z_B}^{\kappa} = g_h$:

$$\frac{g_t}{g_h} = (1-\eta)\frac{{h^*}_t}{h_t} + \eta$$

Where g_t is the growth rate of h, at time t.

We define that at the time of arriving at the island 0, the ratio of the agent's human capital to the equilibrium level of human capital is $\frac{(1+g_k)h_0(z_B,\vartheta)}{(1+\delta_k)h_0(z_A,\vartheta)} = \pi$. Then we have:

$$\frac{g_1}{g_h} = (1-\eta)\pi + \eta$$

$$\frac{g_2}{g_h} = (1-\eta)\frac{\pi}{(1-\eta)\pi+\eta} + \eta$$

$$\frac{g_3}{g_h} = (1-\eta)\frac{\pi}{(1-\eta)\frac{\pi}{(1-\eta)\pi+\eta}+\eta} + \eta$$

… …

Therefore, we can solve the time T when $\frac{g_T}{g_h} = 1$ if $t \geq T$.

As we also know that in steady state $1 + g_k = (1 + g_h)(1 + g_z)$ and $\ln\,(1 + g_z)$ can be cancelled out in the consumption utility gain.

The consumption utility gain from switching island Δu_c is:

$$\left(\frac{(1-\delta_k+r(1-\tau_k)-(1+g_k))}{(1+g_k)}k_0\right) + \sum_{t=0}^{\infty}\beta^{t+1}\ln\Big(\frac{z_B(1-\delta ss_h)(1-\tau_w)}{1+g_h} \times h_0(z_A,\vartheta)F_1(.)(1+g_t)^t + (1-\delta_k+r(1-\tau_k)-(1+g_k)) \times (1+g_h)^t k_0\Big) - \sum_{t=0}^{\infty}\beta^t\ln\Big(z_A(1-\tau_w)h_0(z_A,\vartheta)F_1(.)(1+g_h)^{t-1} + (1-\delta_k+r(1-\tau_k)-(1+g_k))(1+g_h)^t k_0\Big)$$

We are now able to find the locations of agents with different characteristics. We assume that island A is an island with relatively high quality but low technology and that island B is an island with relatively low quality but high technology. Switching from A to B has a positive utility gain from consumption and negative utility gain from quality i.e. $\Delta u_c > 0$, $\Delta u_q < 0$.

Proposition and their testing. **Proposition 1. People with high capital and low speed of learning tend to stay in the island with better environment and people with low capital and high speed of learning tend to stay in the island with higher technology**

The above algebra demonstrates that the consumption gain from island A to B is positively related to the speed of learning. As the utility change of consumption is positively related to the change of consumption, the agent who can learn quickly will receive greater benefit from switching island.

It is also noteworthy that the utility consumption is concave in the consumption, implying that the increase in utility related to the growth of consumption from a low level in likelihood will be more than the utility decrease brought about by switching from a high-quality island to a low-quality island, which is irrelevant to the level of wealth. The reason for this is that the initial level of capital does not affect the change of consumption that is caused by switching island while the initial level of consumption does affect the change of utility. In terms of concave utility, given the fixed change of consumption, the higher the initial level of consumption, the lower the change of utility will be.

Therefore, an agent with low capital that can learn rapidly will be more likely to switch to the island with high technology even though they might have to live in an island with the worse environment. Conversely, agents with high capital that learn slowly will gain less from switching to a high-technology island. They are less like to sacrifice the quality of the environment to earn a higher wage.

Proposition 2. People with high capital and high speed of learning or people with low capital and low speed of learning tend to stay in the island with balanced technology and environment

A high (low) level of capital has a negative (positive) effect on the utility gain while a high (low) speed of learning has a positive (negative) effect. Therefore, agents will balance the benefits of wages and quality and are more likely to locate themselves on an island where the level of technology and capital are relatively proportionate.

Proposition 1 could potentially explain the reason that people in the upper-middle economic class 'escape' from metropolitans. It could also partially explicate the reason for the high concentration of highly educated people and capital in larger cities. However, Proposition 2 alone is not enough to justify the correlation between inequality and urbanisation. Moreover, it also cannot interpret the reasons for labour concentration only occurring happens in particular countries. This suggests that as there are multiple steady-state equilibriums, different assumptions about the distribution of productive, capital and speed of learning could be added to 'narrow' the equilibriums of different countries. Initially, we should assume that the distribution of speed of learning follows a truncated normal distribution in each island from $[\underline{\vartheta}, \overline{\vartheta}]$ and has the same mean and variance, which follows the setting of Pluchino et al. (2018) where the talent of people is randomly assigned. In terms of each level of initial capital owned by the individual, the conditional distribution of the speed of learning to capital follows the same normal distribution as well. We should note that two distributions are independent and it is because the endowment capital is decided by the nature at the time when the city firstly functions in our paper. For example, when migrates find a new continent or they build a new city in a forest, the endowment wealth is the nature resources they have at the time of settling down.

We now start from the economy with two islands A and B and $z_B > z_A$, $q_B < q_A$, which is consistent with what has already been discussed. It is evident that given ϑ, the net gain from switching A to B decreases with capital k. However, for the given k, we find that the part of Δu_c when $t > T$ is:

$$\sum_{t=T}^{\infty}\beta^t\ln\left(\frac{\frac{z_B(1-\delta_h)(1-\tau_w)}{1+g_h}h_0(z_A,\vartheta)F_1(.)\prod_{n=1}^{T-1}(1+g_n)+k_0(1-\delta_k+r(1-\tau_k)-(1+g_k))}{z_A(1-\tau_w)h_0(z_A,\vartheta)F_1(.)(1+g_h)^{T-1}+k_0(1-\delta_k+r(1-\tau_k)-(1+g_k))}\right)$$
$$= \frac{\beta^T}{1-\beta}\ln\left(\frac{\frac{z_B(1-\delta_h)(1-\tau_w)}{1+g_h}h_0(z_B,\vartheta)F_1(.)(1+g_h)^{T-1}+k_0(1-\delta_k+r(1-\tau_k)-(1+g_k))}{z_A(1-\tau_w)h_0(z_A,\vartheta)F_1(.)(1+g_h)^{T-1}+k_0(1-\delta_k+r(1-\tau_k)-(1+g_k))}\right)$$

As we know that $h(z,\vartheta) = (\frac{\vartheta z^\kappa}{g_h})^{\frac{1}{1-\eta}}$, it equals:

$$\frac{\beta^T}{1-\beta}\ln\left(\frac{\frac{z_B(1-\delta_h)(1-\tau_w)}{1+g_h}(\frac{\vartheta {z_B}^\kappa}{g_h})^{\frac{1}{1-\eta}}F_1(.)(1+g_h)^{T-1}+k_0(1-\delta_k+r(1-\tau_k)-(1+g_k))}{z_A(1-\tau_w)(\frac{\vartheta {z_A}^\kappa}{g_h})^{\frac{1}{1-\eta}}F_1(.)(1+g_h)^{T-1}+k_0(1-\delta_k+r(1-\tau_k)-(1+g_k))}\right),$$

which increases with ϑ

Also, as $\frac{(1+g_h)h_0(z_B,\vartheta)}{(1-\delta_h)h_0(z_A,\vartheta)} = \pi$, which is irrelevant of ϑ and, T and g_t are irrelevant ϑ when $t < T$. Therefore, the degree to which the utility of consumption is different at each time before T, which is the same form after T, increases with ϑ. So we can conclude that with the same k, the willingness to move from A to B increases with ϑ.

Based on the analysis above, we can now produce a graph (Fig. 1), which demonstrates visually the location of agents for each k and ϑ.

The curve indicates that given the certain value of k, all agents have a speed of learning, of which the gain and loss from switching A to B is equal and below $\overline{\vartheta}$, the highest limit of talent,

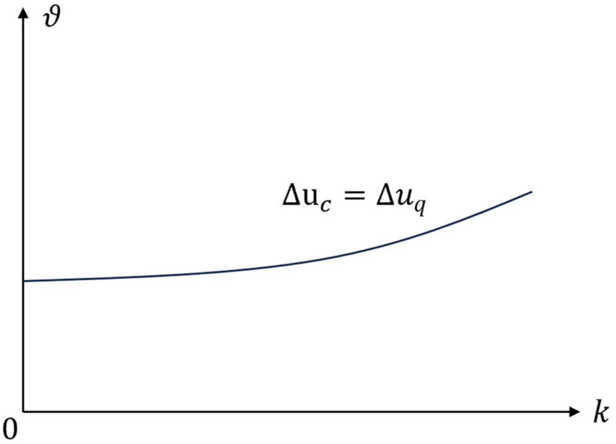

Fig. 1 Choice of the island to stay given their characteristics. For given k, Δu_c requires a higher level of ϑ to be equal to the fixed Δu_q. The shape of curve is convex and $\frac{\partial \Delta u_c}{\partial k \partial \vartheta} > 0$, which implies it needs an increasing marginal value of ϑ for higher level of k to keep the Δu_c constant.

which is not depicted in the graph, but it is assumed that it is above the curve for every k.

As the distribution of ϑ is the same for each k in each island to begin with, we find that the higher the level of k, the fewer agents will be located in the island B, and the higher mean of ϑ for those who are located in the island B. This will lead to income polarisation as all the richest people in island B do not leave whilst the richest people in island A move to island B. This increases the overall number of people that have high wages and capital, and whose income is in the top quartile of income distribution.

With regards to people with lower capital, the lower bound of ϑ to be located in island B decreases, implying that the most impoverish people move to island B. However, as k increases, most people whose ϑ is not high enough will move to island A. These people tend to have an income within the quartile between the richest and poorest, therefore, income inequality becomes more pronounced than in the initial distribution for island B.

The main source of income for people that have a lower capital is their wages and if they do not relocate to the city with higher technological advancement, their consumption will be too low thus they will find it difficult to survive. Therefore, they will often move to chase higher wages even though the level of quality might be lower. The capital income of people who possess a lot capital of is high enough to satisfy their consumption needs; as such they might be more inclined to enjoy the higher quality of life afforded to them by living in a rural and/or seaside city. It is only the richest people whose wages are also high that will opt to stay where they are because they do not want to see a reduction in their income by moving as their wage makes up a significant part of their total income. Furthermore, wage income may not only be a wage given by an employer but can also refer to the income generated by the companies of entrepreneurs which is based in the city where the agent residents.

This could also be the reason that foreign immigrant is seen to gather in metropolitan areas. Migrant workers are generally seen as agents with a low level of capital. The lower bound of ϑ to be in the island B is low, as such most of this group will choose to locate in cities with higher technology. This is consistent with the empirical data gathered in UK where 35% of people who are born outside of the UK live in a capital city (ONS, 2021) and is also consistent with Gordon and Kaplanis' (2014) research which argues that immigrants are the reason for the increase of low paid jobs in metropolitans.

The multiple islands economy will now be examined. No matter what the distribution of the (z, q) of the islands is, one island must

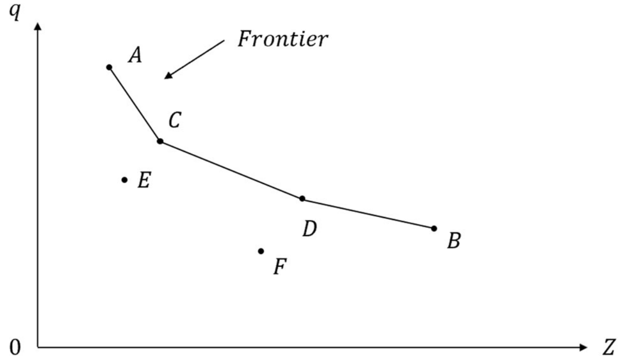

Fig. 2 Pattern of ideal and non-ideal islands. The graph describes the islands with different environments and technologies. The ideal islands should be the ones on the frontier.

be the best in terms of technology and the other in terms of equality. We will call them island B and A for consistence.

For $v_{leave}(k, \vartheta)$ in the steady state of every level of k and ϑ, it searches for islands with z, q to maximise

$$\ln\left(\frac{\left(1-\delta_k+r\left(1-\tau_k\right)-\left(1+g_k\right)\right)}{1+g_k}k_0\underline{q}\right)$$
$$+\sum_{t=0}^{\infty}\beta^{t+1}\ln\left(\frac{z\left(1-\delta_h\right)\left(1-\tau_w\right)}{1+g_h}h_0(z,\vartheta)F_1(.)\left(1+g_z\right)^t\left(1+g_t\right)^t\right.$$
$$\left.+qk_0\left(1+g_k\right)^{2t+1}\left(1-\delta_k+r\left(1-\tau_k\right)-\left(1+g_k\right)\right)\right)$$

When desiring to move, agents will be able to find their corresponding ideal island and if we draw all the (z,q) sets on the 2-dimensional graph, the necessary condition for ideal islands is that they must lie on the northeast frontier (connecting all of the points, which has no points on the northeast side in Fig. 2). In terms of Island E and F, they cannot be ideal islands as there are islands with both a higher z and q.

When comparing B and D, the lower bound of $\vartheta_B(k)$ can be found above which the agent prefers B to D and below which the agent prefers B to D. Then we compare D and C and trace the lower bound of $\vartheta_D(k)$ above which the agent prefers C to D and below which the agent prefers C to D. Again, going through the same steps, we can outline $\vartheta_C(k)$ above which the agent prefers C to A and below which the agent prefers C to A. Therefore, as the agents are rational, which implies the transitivity of their preference, the ideal island of the agent with ϑ and k in the region $[\vartheta_x(k), \vartheta_y(k)]$ is y. Moreover, as we know that the lower bound $\vartheta_y(k)$ decreases with k, it is possible to draw the graph (Fig. 3) describing the ideal location of agents with different ϑ and k:

Agents that will move are now the focus of the research. As has been established, on the frontier, a higher quality island must be accompanied with lower levels technology. For each level of k and ϑ, the conclusion of the two islands economy for each two neighbour islands can be applied to find the ideal island for each agent. The agent with highest level of k and the island with highest level and second highest level of z is the starting point. For simplicity, we continue to use the six islands case (it can be applied to any number of islands with the only difference being the times of the comparison).

Given k, as an agent's choice of an ideal island is determined solely by the speed at which the agent learns, their income is determined by it as well. This means that the analysis of the income distribution can be simplified to the speed of learning distribution.

However, for each region of y on each island, not all agents will move to their ideal island as the cost of moving and living there might be too high for them to afford i.e. $\Delta u_c + \Delta u_q < 0$. For each

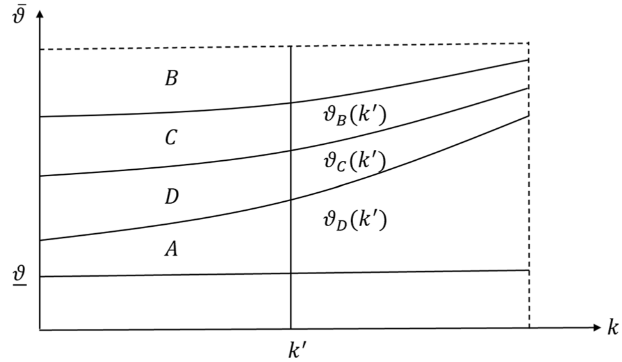

Fig. 3 Region in which the agent with k and ϑ will find the ideal island. The areas in the graph are the ideal islands to settle down for the individuals with corresponding knowledge and wealth.

island i, the threshold value of ϑ_y^i can be calculated, where the agent is indifferent about whether they will move to island y or not ($y \neq i$).

In other words, the Fig. 3 is different from Fig. 1, as Fig. 3 indicates the island where the agent wants to move if the agent is forced to be 'on the way' to search for an island for one period and arrive to a island at the next period. This island is called 'ideal' island. Therefore, there are agents who want to stay at the current islands if they are not forced to move. ϑ_y^i is the threshold of talents of people in island i who are indifferent between moving to island y. $\vartheta_y^i(k)$ is actually the $\vartheta(k)$ of Fig. 1.

For the neighbouring island i, $z_i > z_y = > \vartheta_y^i < \vartheta_y$ and $z_i < z_y = > \vartheta_y^i > \vartheta_y$.

This is because in terms of the original island and the target island, the absolute value of Δu_c is positively correlated with ϑ. Moreover, moving to other islands incurs additional costs of utility; moving to islands with higher technology requires higher ϑ to gain more income and offset the costs. With the same logic, moving to a lower technology island will require lower ϑ to experience a smaller decrease in income. We also assume that ϑ_y^i is within the region y to ensure that there are agents moving to all other islands.

All the people in region B stay in island B and some of people from region B in other islands move to island B. Furthermore, there is an outflow of people from island B to other regions, which increases the proportion of region B people who are the richest on island B. With the same logic, island A faces a concentration of the poorest people as well as an outflow of people in other regions. In both island A and B, inequality is increased compared to their original states.

However, regarding island C and D, we see an outflow of both the richest and poorest people and inflow of middle-class people, which causes inequality to decrease.

The unattractive islands like E and F face a population outflow and only those who suffer a loss of utility because of bad quality and the loss of wages during moving which offsets the gain of moving to the ideal island will stay. The percentage of the population that stays is non-decreasing with z and q given the distribution of (k, ϑ) fixed.

The conclusion we can draw from that is that, for given k, wage inequality decreases then increases as the technology of the island also advances. Therefore, as the wage distribution starts out as the same (truncated normal), the total wage inequality also follows the same trend as the one conditional on k.

As k decreases, the area of region B increases, and the concentration of region B agents moving from other islands to island B also goes up. This leads to an ambiguous effect on inequality as it did for the high-technology island. For each k (i) the average income decreases as k decreases but (ii) the proportion of agents with capital k to the total number of agents in that island increases. It is impossible to know which effect dominant, therefore we also cannot know why the income distribution has shifted. The effects are the opposite for island A compared to those of B as k increases, but still the two effects conflict with each other.

Extended two-goods model

Description. The difference compared to the baseline model is that we now have two consumption goods:

A. Tradable goods, which can be traded between islands, must have the same price for all islands, as such we can normalise the price as 1. They will be referred to as good A.
B. The non-tradable goods, which can only be consumed locally, will have different prices depending on the island. They will be called good B.

The utility of the agent now is

$$\mathrm{u}(\mathrm{c_t}, \mathrm{q_t}) = \ln(\mathrm{c_{At}}) + \ln(\mathrm{C_{Bt}}) + \ln(\mathrm{q_t})$$

We now have Bellman equation for the agent:

$$\nu(\omega, z, h, \mu, q, k, \vartheta, K) = \max\{\nu_{stay}(\omega, z, h, \mu, q, k, \vartheta, K), \nu_{leave}(k, h, \vartheta, \mu)\}$$

where

$$\begin{aligned}&\nu_{stay}(\omega, z, h, \mu, q, k, \vartheta, K)\\ &= \max_{k'}\{u(h(1-\tau_w)w(\omega, z, \mu, q, k) + r(\mu)(1-\tau_k)k + (1-\delta_k)k - k'\\ &\quad - p_B(\omega, z, \mu, q, K)c_B, c_B, q)\\ &\quad + \beta E_t \nu(\omega'(\omega, z, \mu, q, K, \vartheta), z'(z, \omega, q, K), h'(z, h), \mu'(\mu), k', q'(q, \mu))\}\end{aligned}$$

$$\begin{aligned}&\nu_{leave}(k, h, \vartheta, \mu)\\ &= \max_{k'}\{u(r(\mu)(1-\tau_k)k + (1-\delta_k)k - k' - p_B(\tilde{\omega}, \tilde{z}, \tilde{q}, \mu, K)c_B, c_B, \underline{q})\\ &\quad + \beta \max_{\tilde{\omega},\tilde{q},\tilde{z},\tilde{K}}\{E_t \nu(\omega'(\tilde{\omega}, \tilde{q}, \tilde{z}, \tilde{K}, \mu), z'(\tilde{\omega}, \tilde{q}, \tilde{z}, \tilde{K}), h(1-\delta_h), \mu'(\mu), k', q'(\tilde{q}, \mu), \vartheta)\}\}\end{aligned}$$

The modification is that an agent can immediately arrive at any of the other islands without delay but can only start working during the next period.

z_A and h follow the growth path $\dot{z_A} = \theta(H_A)^\lambda {z_A}^\phi {z_n}^\gamma$ and $\dot{h} = \vartheta h^\eta {z_A}^\kappa$ and H_A is the human capital working to produce A. Following the deviation performed in part I, we have in steady state:

$$g_z = \frac{1-\eta}{(1-\phi)(1-\eta)-\lambda\kappa} g \text{ and } g_h = \frac{\kappa}{(1-\phi)(1-\eta)-\lambda\kappa} g$$

The technology of the non-tradable good z_B is the same for all islands as it grows at the constant rate: $g_{zB} = g_z = \frac{1-\eta}{(1-\phi)(1-\eta)-\lambda\kappa} g$, i.e. all technologies grow at the same rate in steady state.

There are two production sectors in each island for two goods. The sector of good A is the same, which is homogeneous with degree 1, to solve:

Given $w(\omega, z_A, \mu, q)$ and $r(\mu)$

$$\max_{K_A, H_A}\{F_A(z_A H_A, K_A) - w(\omega, z_A, \mu, q)H_A - r(\mu)K_A\}$$

The sector of good B, with its price $p_B(\omega, z_A, \mu, q)$ given, is also homogeneous with degree 1 and it solves:

$$\max_{K_B, H_B}\{p_B(\omega, z_A, \mu, q)F_B(z_B H_B, K_B) - w(\omega, z_A, \mu, q)H_B - r(\mu)K_B\}$$

For consistence, in each island: $H_A + H_B = \int h * e(k, h, \vartheta | z_A, q, \omega) dk dh d\vartheta$ and $K_A + K_B = K$

Equilibrium. We solve the first conditions for two sectors and apply the law of one price on Appendix 2.

We now can solve other conditions for steady state:

Condition 1 – Agents with initial value of k_0,h_0,ϑ at the time of steady state, will only be on the island with

$$z_{A0}, q_0 \in \left\{ z_{A0}, q_0 | v_{stay}(z_{A0}, q_0, k_0, h_0, \vartheta) \geq v_{leave}(k_0, h_0, \vartheta) \right\}$$

Condition 2 – The sum of capital for each island should equal the sum of capital for each agent:

$$\int k\omega(dk, d\vartheta | q, K, z_A \mu(dq, dK, dz_A) = \int K\mu(dq, dK, dz_A)$$

Condition 3 – With the same logic, for each island i, the human capital employed by the company should be equal to the sum of human capital in that island:

$$\int h(\vartheta, z_A)\omega(d\vartheta, d\kappa | q, K, z_A) = H_A(z_A) + H_B(K, z_A)$$

Condition 4 – At each island the total production of non-tradable goods is equal to their total consumption:

$$\int c_B(\vartheta, k, z_A)\omega(dk, d\vartheta | q, K, z_A) = F_B(z_B H_B(z_A, K), K_B(z_A, K))$$

Similarly to part I, any distribution of ω and μ as well as satisfying the above conditions can suffice as the equilibrium, as such there are multiple equilibriums.

Compared with part I, the locations of labour based on the characteristics and the differing distributions of islands are comparable while people's willingness to move to the higher quality cities is determined by the lower prices-wage ratio of their non-tradable goods. The region of area B and the concentration to island B both perform a decrease. Therefore, with regard inequality, island B as a lower rate and island A has a higher rate when compared to part I.

Adding the non-tradable good improves the previous model by imposing stricter conditions for equilibrium. This suggests that in the islands with higher technology, the non-tradable goods have higher price even though these goods are the same for all islands. The intuition is that higher technology islands have higher wage for the tradable goods. To ensure that people willing to work for the non-tradable sector, two sectors must have the same wage. To cover the additional cost from the wage, the non-tradable sectors must increase the price. This is the Balassa–Samuelson effect (Samuelson, 1964) and Kravis and Lipsey (1983), which is found statistically significant by Samuelson (1994). Vaona (2011) and Nenna (2001) provide evidence supporting it from intra-country level in Italy, and Songtao (2009) tests it for real estate's price between cities in China, indicating that the cities' higher technology level is positively correlated with the wage and price of real estate sector, which is the rationale of applying Balassa–Samuelson effect on the city-level.

The adjusted model used in this section leads to a similar distribution when contrasted with that of part I. However, the non-tradable goods provide another intuition to explain labour mobility and its causes. The labours also consider the price of non-tradable goods as a factor determining the decisions of labour mobility. When considering where to work, people do not only consider the wage but also the purchase power of it on local food, service, or accommodation. For the labours with higher talent but lower wealth, they mainly rely on wage income. As we discussed in the Appendix 2, the wage in large cities is more affordable for non-tradable goods. Comparing to proposition 1 of part 2, they are more likely to move to the cities with higher technology. Conversely, for labours with lower talent but higher wealth who mainly rely on capital income, they are more likely to move to the cities with better environment because the price level is higher in large cities. For people with low talent and low wealth or high talent and high wealth, the conclusion of proposition 2 of part 2 is the similar as they also balance between the gain of more affordable wage income and less affordable capital income from moving to a higher technology city.

It also accounts for the low wage-price ratio in cities with superior environments. This is consistent with the positive correlation between the natural environment and house prices in previous literature. Luttik (2000) and Donovan et al. (2019) estimate the effect of demographic factors on house prices and reveal the significant positive correlation between trees, water, open spaces, and house prices. Furthermore, Catte et al. (2004) and Steegmans and Hassink (2017) find that wealth of the buyers has a significantly positive effect on housing prices. Generally, housing prices are also positively correlated to the level of local wages. However, there are also cities with superior environment where local wages are relatively low also with high housing prices. Therefore, we might assume that areas with higher house prices are more likely to be those with higher-income residents whose wealth drives up house prices. This implies that cities with relatively low technology but a relatively high quality of environment might also attract households with high incomes. Achieving a high income in the city with low wages is brought about through these households that have capital income from other cities. This implies that if a city government spends too much on developing tourism while ignoring the investment on R&D, the population in that city will be people with high capital store and low skill, which drive up the price level but do not help with the wage increase. A typical city in China is Sanya, which is famous for its natural environment. Its house price-wage ratio is of top 5 in China while its average wage is even not in top 50 (Chinese National Bureau of Statistic, 2022).

Empirical analysis

We apply the empirical analysis to support the Lucas-Prescott style island model. The empirical analysis is based on the American population movement data in IPUMS USA from 2016–2021 at state level. And we use the patent data from US patent and trademark office to proxy for the technology level, and the higher number of patents indicates the high level of state technology level. The air quality index (AQI), the comprehensive evaluation of pollutants such as SO2, NO2, PM10, PM2.5, O3 and CO, from US Environmental Protection Agency to measure the environmental level. While, the bigger AQI implies the worse state environment.

To test the impact of individual wealth and education on the living choice, we construct the following Logit model:

$$\ln \frac{\text{Probability}(\text{Environmental}_{i,j,t}=1)}{\text{Probability}(\text{Environmental}_{i,j,t}=0)} = \alpha_0 + \beta_1 Education_{i,j,t} + \beta_2 Wealth_{i,j,t} + Controls + Year + WorkInd + County \quad (1)$$

$$\ln \frac{\text{Probability}(\text{Technology}_{i,j,t}=1)}{\text{Probability}\left(\text{Technology}_{i,j,t}=0\right)} = \alpha_0 + \beta_1 Education_{i,j,t} + \beta_2 Wealth_{i,j,t} + Controls + Year + WorkInd + County \quad (2)$$

and Probit regression model:

$$\text{Probability}(Environmental_{i,j,t} = 1) = \Phi(\alpha_0 + \beta_1 Education_{i,j,t} + \beta_2 Wealth_{i,j,t} + Controls + Year + WorkInd + County) \quad (3)$$

$$\text{Probability}(Technology_{i,j,t} = 1) = \Phi(\alpha_0 + \beta_1 Education_{i,j,t} + \beta_2 Wealth_{i,j,t} + Controls + Year + WorkInd + County) \quad (4)$$

where i represents citizen, j represents state and t represents year. *Environmental* takes value of 1 if individual i lives in a state j with good environment quality and 0 otherwise, by the median of AQI. And Tech*nology* takes value of 1 if individual i lives in a state j with high technology and 0 otherwise, by the

Table 1 Empirical analysis.

Variable	(1) Logit Odds ratio *Environment*	(2) Logit β *Environment*	(3) Logit Odds ratio *Technology*	(4) Logit β *Technology*	(5) Probit Odds ratio *Environment*	(6) Probit β *Environment*	(7) Probit Odds ratio *Technology*	(8) Probit β *Technology*
Education	**0.978*****	**−0.022*****	**1.005*****	**0.005*****	**0.987*****	**−0.013*****	**1.003*****	**0.003*****
	(0.001)	**(0.001)**	**(0.001)**	**(0.001)**	**(0.001)**	**(0.001)**	**(0.001)**	**(0.001)**
Wealth	**1.015*****	**0.015*****	**0.990*****	**−0.010*****	**1.009*****	**0.009*****	**0.994*****	**−0.006*****
	(0.002)	**(0.002)**	**(0.002)**	**(0.002)**	**(0.001)**	**(0.001)**	**(0.001)**	**(0.001)**
Sex	1.001	0.001	1.006	0.006	1.001	0.001	1.004	0.004
	(0.005)	(0.005)	(0.005)	(0.005)	(0.003)	(0.003)	(0.003)	(0.003)
Age	1.002***	0.002***	1.001	0.001	1.001***	0.001***	1.001	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Income	1.006***	0.006***	1.005***	0.005***	1.004***	0.004***	1.003***	0.003***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
Farm	0.988	−0.012	0.930***	−0.073***	0.991	−0.010	0.957	−0.044
	(0.075)	(0.076)	(0.071)	(0.076)	(0.045)	(0.045)	(0.045)	(0.045)
Constant	0.761***	−0.273***	1.155**	0.144**	1.186***	−0.171***	0.912***	0.092***
	(0.037)	(0.048)	(0.056)	(0.048)	(0.029)	(0.029)	(0.028)	(0.028)
Observations	734,440	734,440	734,440	734,440	734,440	734,440	734,440	734,440
Pseudo R2	0.1982	0.1982	0.2156	0.2156	0.1982	0.1982	0.2155	0.2155
Year FE	YES	YES	YES	YES	YES	YES	YES	YES
WorkInd FE	YES	YES	YES	YES	YES	YES	YES	YES
County FE	YES	YES	YES	YES	YES	YES	YES	YES

Robust standard errors in brackets. *significant at the 10 percent level; **significant at the 5 percent level; ***significant at the 1 percent level.
The bold values is the focus as the effect of individual education and wealth level on movement choice.

median of patent number. Φ is the cumulative distribution function of the standard normal distribution. Education level (Education) is the proxy variable for individual skill, measured as the sequence from 1 to 11 representing the lowest education level to the highest level and the value of home with logarithm is used to proxy for individual wealth. Controls is a vector of individual control variables include sex, age, logarithm of one plus income (including rental income from housing) and owning farm or not. And the descriptive statistics of all variables are reported in Appendix 3. To control unobservable factors potentially affecting living choice, we further control the year, work industry and county fixed effects, with standard error clustered at family level.

Following Taiwo (2013), the results are presented Table 1, with Columns (1)–(4) reporting Logit model and Columns (5)–(8) reporting Probit model. And columns (1), (3), (5) and (7) report the result of odds ratio, with others reporting the coefficients. The Columns (1)-(2) and (5)-(6) show that the probability of citizens living in states with poor environment increases with their skill level and decreases with their wealth. And Columns (3)-(4) and (7)-(8) reveal that those owning little wealth with high skill level have higher chance to live in state with relatively high technology level. In conclusion, citizens with high skill-level and little wealth tend to choose the state with poor environment and high technology level, and these with low skill-level and huge wealth prefer to live in state with good environment and low technology level. This is consistent with the pervious theoretical analysis that, as higher technology of the location and education level of the individual bring higher wage level, those with little endowment wealth but higher education level have to work in places with higher technology but worse environment to enjoy higher wage to cover their expense. On the contrast, labours with higher wealth tend to live in a place with better environment as they can benefit from capital rent wherever they live and they can gain more utility from the environment. However, we only regress the labour choice at the state level. The further research will be conducted to test if the theory still holds at the county/city level, which is more precise and convincible.

Conclusion

This paper describes and derives a Lucas-Prescott style island model, to study the choices of heterogeneous agents' decision about where they want to live. Differing from other islands model which assume the implementation of the stochastic technology process, it utilises endogenous technology growth, which in turn influences personal human capital growth. This leads to the balanced growth equilibrium in steady state, which is potentially the first model that combines balanced growth and the Lucas-Prescott model. Compared to the conventional Lucas-Prescott models where only the distribution is fixed in steady state, this model provides the fixed location of each agent, which admittedly makes it less flexible. However, that is compensated by less restrictive steady state conditions, which leads to multiple steady states.

The baseline model's steady state indicates that the islands (cities) where people ideally want to stay in must be the islands on the frontiers (see Fig. 2). Among these islands, agents that have a higher speed of learning and lower capital tend to reside in urban metropolitans, while people who rely mainly on non-wage-based income prefer to live in a city with better environment. It leads to the U-shape curve of the wage income inequality in terms of the technology of these islands, which is consistent with Nord's (1980) research. However, in terms of total income, the inequality is ambiguous. The two-goods model includes both tradable and non-tradable goods, which interprets the high price-wage ratio of the cities with good environments through the Balassa–Samuelson effect. Even though the implications behind the choice of a location still holds, its magnitude is increased by the impact the price of non-tradable goods.

However, this paper does not consider the role of the government to allocate resources to attract workers, support technology development, and improve the environment. It also lacks the deep analysis of the real-world data. To complete the research, the future study will add the government to the model and suggest different policies for the government with different goals. Then we will simulate the model, which is calibrated based on the data of US. Also, we will evaluate the model based on microdata of individual labour at the county/city level. For example, apart from the labour choice, which is mentioned before, macro-level data of each city's price level will be regressed to examine the fit of the simultaneous effects of migration and the price of non-tradable goods on each other to the model.

Data availability

The data generated during and/or analysed during the current study are available in the supplementary files.

Received: 6 November 2023; Accepted: 19 August 2024;
Published online: 05 September 2024

Notes

1 The level of urbanization is calculated as the proportion of citizens in the urban area to the total population within the country. The data is accessible from World Bank database at: https://data.worldbank.org/indicator/SP.URB.TOTL.IN.ZS
2 The data is accessible from World Bank database at: https://data.worldbank.org/indicator/EN.URB.LCTY.UR.ZS
3 Metropolitan Statistical Areas are core based statistical areas (CBSAs) associated with at least one Urban Area that has a population of at least 50,000. The metropolitan statistical area comprises the central county or counties or equivalent entities containing the core, plus adjacent outlying counties having a high degree of social and economic integration with the central county or counties as measured through commuting, according to the definition of US Census at: https://www.census.gov/programs-surveys/geography/about/glossary.html#par_textimage_7.

References

Abel JR, Dey I, Gabe TM (2012) Productivity and the density of human capital. J Reg Sci 52(4):562–586
Aiyagari SR (1994) Uninsured idiosyncratic risk and aggregate saving. Q J Econ 109(3):659–684
Aksoy CG, Poutvaara P (2021) Refugees' and irregular migrants' self-selection into Europe. J Dev Econ 152:102681
Bargain O, Kwenda P (2011) Earnings structures, informal employment, and self-employment: new evidence from Brazil, Mexico, and South Africa. Rev Income Wealth 57:S100–S122
Baum-Snow, N and Pavan R (2009a) Inequality and city size. Brown University, US, Rhode Island
Behrens K, Robert-Nicoud F (2014) Survival of the fittest in cities: Urbanisation and inequality. Econ J 124(581):1371–1400
Borjas GJ (1987) Self-selection and the earnings of immigrants. Am Econ Rev 77(4):531–553
Buch T, Hamann S, Niebuhr A, Rossen A (2014) Attractiveness of regions and cities for talents and entrepreneurs. Econ Geogr 90(Supplement S1):59–77
Cao G, Li M, Ma Y, Tao R (2015) Self-employment and intention of permanent urban settlement: evidence from a survey of migrants in China's four major urbanising areas. Urban Stud 52(4):639–664
Catte, P, Girouard, N, Price, R, & André, C (2004). Housing markets, wealth and the business cycle
Castillo V, Figal Garone L, Maffioli A, Rojo S, Stucchi R (2020) Knowledge spillovers through labour mobility: an employer–employee analysis. J Dev Stud 56(3):469–488
Chen S, Oliva P (2022) The effect of air pollution on migration: evidence from China. China. J Dev Econ 156:102833
Chin, A, & Cortes, KE (2015). The refugee/asylum seeker. In: Handbook of the economics of international migration. 585–658. North-Holland, Elsevier
Chinese National Bureau of Statistic. (2022). Statistical Yearbook, various issues
Coen-Pirani D (2010) Understanding gross worker flows across US states. J Monet Econ 57(7):769–784
Donovan GH, Landry S, Winter C (2019) Urban trees, house price, and redevelopment pressure in Tampa, Florida. Urban For Urban Green 38:330–336
Feld B, Nagy A, Osman A (2022) What do jobseekers want? Comparing methods to estimate reservation wages and the value of job attributes. J Dev Econ 159:102978
Glaeser EL, Resseger M, Tobio K (2009) Inequality in cities. J Reg Sci 49:617–646
Gordon IR, Kaplanis I (2014) Accounting for big-city growth in low-paid occupations: immigration and/or service-class consumption. Econ Geogr 90(1):67–90
Gong X, Van Soest A (2002) Wage differentials and mobility in the urban labour market: a panel data analysis for Mexico. Labour Econ 9(4):513–529
Harris JR (1970) Migration, unemployment and development: a two-sector analysis. Am Econ Rev 60(1):126
Hopenhayn HA (1992) Entry, exit, and firm dynamics in long run equilibrium. Econom: J Econom Soc 60(5):1127–1150
Kravis IB, Lipsey RE (1983). Toward an explanation of national price levels, vol 52. International Finance Section, Department of Economics, Princeton University
Krusell P, Smith Jr AA (1998) Income and wealth heterogeneity in the macroeconomy. J Polit Economy 106(5):867–896
Lucas Jr, RE, & Prescott, EC (1978). Equilibrium search and unemployment. In Uncertainty in economics. Academic Press, UK, Cambridge, Massachusetts, pp 515–540
Luttik J (2000) The value of trees, water and open space as reflected by house prices in the Netherlands. Landsc Urban Plan 48(3-4):161–167
Madden, JF (2000) Changes in Income Inequality within U.S. Metropolitan Areas. Kalamazoo, MI: W.E. Upjohn Institute for Employment Research
Ministry of Housing, Communities & Local Government. (2020). Land Use in England, 2018
McCall JJ (1970) Economics of information and job search. Q J Econ 84(1):113–126
Nenna, M (2001). Price level convergence among Italian cities: any role for the Harrod-Balassa-Samuelson hypothesis? Universita degli studi di Roma" La Sapienza"
Nord S (1980) An empirical analysis of income inequality and city size. South Econ J 46(3):863–872
Office for National Statistics. (2021). House price to residence-based earnings ratio. https://www.ons.gov.uk/peoplepopulationandcommunity/housing/datasets/ratioofhousepricetoresidencebasedearningslowerquartileandmedian
Pluchino, A, Biondo, A, Rapisarda, A (2018). Talent vs Luck: the role of randomness in success and failure
Romer PM (1986) Increasing returns and long-run growth. J Polit Economy 94:1002–1037
Samuelson, PA (1964). Theoretical notes on trade problems. The review of economics and statistics, 145-154
Samuelson PA (1994) Facets of Balassa-Samuelson thirty years later. Rev Int Econ 2(3):201–226
Songtao, W (2009). Urban Openness and Real Estate Prices: Empirical Evi-dence from Thirty—Five Large Scale Chinese Cities [J]. Nankai Economic Studies, 2
Steegmans J, Hassink W (2017) Financial position and house price determination: an empirical study of income and wealth effects. J Hous Econ 36:8–24
Taiwo O (2013) Employment choice and mobility in multi-sector labour markets: theoretical model and evidence from Ghana. Int Labour Rev 152(3-4):469–492
Van Nieuwerburgh S, Weill PO (2010) Why has house price dispersion gone up? Rev Econ Stud 77(4):1567–1606
Vaona A (2011) Intra-national purchasing power parity and Balassa–Samuelson effects in Italy. Spatial Econ Anal 6(3):291–309
Wheeler CH (2001) Search, sorting, and urban agglomeration. J Labor Econ 19:879–899
Zenou Y (2008) Job search and mobility in developing countries. Theory and policy implications. J Dev Econ 86(2):336–355
Zhao X(2020) Land and labor allocation under communal tenure: theory and evidence from China. J Dev Econ 147:102526

Author contributions

Author contributions are as follows: Conceptualization, T.Q.; methodology, T.Q.; software, Y.H.; validation, Y.G.; formal analysis, TQ. and Y.H.; investigation, Y.G.; resources, Y.H.; data curation, Y.H.; writing—original draft preparation, T.Q., Y.H. and Y.G.; writing—review and editing, Y.H; visualization, Y.H.; supervision, T.Q. and Y.G.

Competing interests

The authors declare no competing interests. None of the authors on the paper are members of this journal's editorial board.

Ethical approval

Ethical approval was not required as the study did not involve human participants. The procedures used in this study adhere to the tenets of the Declaration of Helsinki.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1057/s41599-024-03627-9.

Correspondence and requests for materials should be addressed to Yongjian Huang.