CIRJE-F-496

Estimating Interregional Utility Differentials

Kentaro Nakajima
Graduate School of Economics, University of Tokyo

Takatoshi Tabuchi
University of Tokyo

May 2007

Estimating Interregional Utility Differentials*

Kentaro Nakajima† Takatoshi Tabuchi‡

29 May 2007

Abstract

The examination of long-term Japanese data on interregional migration revealed three stylized facts of migration behavior. Based on the facts, we formulated an operational model and estimated interregional utility differentials. We found that the interregional utility differentials have been converging until the late 1970s. We showed that the utility estimates are highly correlated with per capita real income. We also applied the model to interregional migration in the United States and Canada as well as the interindustry movement in Japan and confirmed the model's validity.

Keywords: migration costs; payoff monotonicity; revealed preference; gross migration

J.E.L. Classification: R11, R23

1 Introduction

There are diverse reasons for human migration. These reasons are not just based on economic factors but also on non-economic factors. In the process of urbanization and economic development, one-way migration from rural to urban areas is dominant. This is because economic factors such as wage differential and employment opportunities are crucial. In developed countries, however, the amount of net migration has decreased significantly compared to the gross migration recently. This may indicate the importance of non-economic factors such as marriage, admission into school, retirement, and regional amenities. Thus, we should consider utility in a broader sense such that it includes both economic and non-economic factors.

It is widely observed in developed countries that there are many migration flows from regions i to j, and there are many migration flows from j to i. The prevalence of such two-way migration may imply that migration is based on non-economic factors, which differ among individuals. On

*Thanks go to Y. Kasajima for excellent research assistance.

†Graduate School of Economics, University of Tokyo, nakajima.kentaro@gmail.com

‡Faculty of Economics, University of Tokyo, ttabuchi@e.u-tokyo.ac.jp

the other hand, net migration flows are much smaller than gross ones. This may indicate that economic factors are losing their significance nowadays.

Crozet (2004) and Pons, Paluzie, Silvestre and Tirado (2007) estimated interregional migration flows based on the new economic geography (Tabuchi and Thisse, 2002), wherein migrants consider real income, i.e., economic factors when making migration decisions. However, they would also take into account non-economic factors such as regional amenities. Kahn (1995) and Cragg and Kahn (1997) applied the hedonic approach by estimating regional wages and rent, which are considered to capitalize differentials in regional utilities including amenities.[1]Insofar as utilities constitute such a large number of factors, it may be difficult to incorporate all of the factors when estimating the utilities. Therefore, we instead estimate the utility values directly by using migration data. This is based on Samuelson's (1948) revealed preference as well as Tiebout's (1956) voting with one's feet.

In this paper, we focus on measuring the interregional utility differentials in the presence of migration costs. In the next section, we present empirical evidence of the stylized facts revealed in migration studies. Based on these stylized facts, we construct an operational model for utility estimation in section 3. Its robustness is checked with regard to various aspects in section 4. Section 5 concludes.

2 Three stylized facts on migration

People do not migrate very frequently. According to Japanese statistics, the average ratio of the annual gross migration to the population is 2.9%. Since the average Japanese life span is 81 years, people engage in interprefectural migration only 2.3 times in their entire life. However, this figure includes residential relocation without changing one's job, and hence, people migrate across regions only twice in their entire life. The very small ratio of migrants to non-migrants is attributed to the high fixed costs of interregional migration, which also includes non-pecuniary costs like local information and human relations. Therefore, *the first stylized fact is the existence of migration costs.*

Net migration is less than gross migration. Denote the gross migration from regions i to j for a given period of time by m_{ij}, and the corresponding net migration by $\Delta m_{ij} \equiv m_{ij} - m_{ji}$. Let n be the number of regions in the space-economy. The interprefectural gross migration $\sum_{i=1}^{n} \sum_{j=1,i\neq j}^{n} m_{ij}$ was 166 million during the last 52 years, and there is no decreasing trend of gross migration as seen in Figure 1. On the other hand, the interprefectural net migration

[1]See Cushing and Poot (2004) for extensive reviews on the recent migration literature.

$\sum_{i=1}^{n}\sum_{j=1,i<j}^{n}|\Delta m_{ij}|$ was 15.9 million during the same period, and has been decreasing over time. Since the ratio of net migration to gross migration is only 9.6%, it can be stated that more than 90% of the migration is "seemingly wasteful".[2][3] Thus, *the second stylized fact is seemingly wasteful migration*, which, however, must be rational once we introduce individual heterogeneity, as will be explained below.

Rational individuals would migrate from low- to high-utility regions although the speed of migration adjustment is not as fast as that of price adjustments. In evolutionary game theory (Weibull, 1995), this property is referred to as payoff monotonicity, wherein a region with a higher utility has a higher growth rate of population. In other words, the direction of net migration coincides with the utility differentials:

$$\operatorname{sgn}(\Delta m_{ij}) = \operatorname{sgn}(U_j - U_i) \tag{1}$$

where U_j is the intertemporal utility in region j defined by the discounted stream of instantaneous utilities in each period. If the payoff monotonicity (1) holds for any pair of regions, migration flows are said to be transitive:

$$\text{if } \Delta m_{ij} > 0 \text{ and } \Delta m_{jk} > 0\text{, then } \Delta m_{ik} > 0 \tag{2}$$

due to transitivity of the utilities: if $U_i < U_j$ and $U_j < U_k$, then $U_i < U_k$. Stated differently,collective preferences are not cyclic, which is in agreement with Tiebout's voting with one's feet.

Since there were $n = 46$ prefectures until 1972 and $n = 47$ prefectures from 1973 in Japan, there are ${}_nC_3$ combinations of net migration among three regions, and ${}_nC_3 \times 52 \approx 800,000$ combinations of net migrations among the three regions for 52 years. Computing all combinations, we found that 88% of them satisfy the transitivity condition (2). If the prefectures are aggregated into 9 regions as will be done later, then 95% of them satisfy the transitivity

[2] This 90% refers to the Grubel-Lloyd (1971) index in international trade, which is defined by

$$GL = 1 - \frac{\sum_i |X_i - M_i|}{\sum_i (X_i + M_i)}$$

where X_i and M_i are exports and imports in region i. If X_i and M_i are the sums of out-migration and in-migration respectively, we have

$$GL = 1 - \frac{\sum_i \sum_{j \neq i} |\Delta m_{ij}|}{2\sum_i \sum_{j \neq i} m_{ij}} = 0.9.$$

[3] These values in the United States during the 16 years from fiscal year 1989 to 2004 are as follows. The average ratio of the annual gross migration to the population is 1.1%, the interstate gross migration is 45.7 million the interstate net migration is 6.4 million, and the ratio of net migration to gross migration is 14%. Hence, 86% of the migration is "seemingly wasteful".

condition.[4] Hence, the volume of anti-transitive migration is extremely small compared to that of transitive migration, evidencing Tiebout's voting with one's feet hypothesis. We can therefore state that *the third stylized fact is payoff monotonicity.* Individuals tend to migrate from low- to high-utility regions, whose rankings are in agreement among heterogeneous individuals at a macro level. It thus follows that the utility ranking of regions may be possible based on the migration data.

3 Analytical framework for utility estimation

As mentioned in the introduction, there are short-run adjustment of commodity markets and long-run adjustments in the labor markets. Excess demand for goods and services is quickly diminished by interregional movements of goods and interregional adjustments in prices, wages and rent in the short run. However, the adjustments in the interregional labor markets take time due to diverse constraints. For example, some people do not migrate immediately because they are too young to enter interregional job markets, because they are bound by a multiyear contract at their firms, or because their psychic costs of adapting to new environments are considerable. In this paper, we focus on the long-run adjustments by paying attention to the above mentioned three stylized facts, and formulate a model that satisfies the following criteria.

(i) As per the first stylized fact, interregional migration costs should be explicitly incorporated so that the percentage of non-migrants is considerable.

(ii) As per the second stylized fact, a microeconomic foundation, not for net migration, but for gross migration should be established.

(iii) As per the third stylized fact, the model should satisfy the payoff monotonicity (1).

The first and second stylized facts may be captured by the introduction of heterogeneity in the individual preference of regions. This may be possible by discrete choice models, such as logit and probit (Anderson, de Palma and Thisse, 1992), where individual preference in their perception of the attributes and characteristics associated with a particular region is heterogeneous. These models have a common microfoundation that each individual maximizes one's utility under imperfect information or under heterogeneity in the following manner. An individual residing in i will decide to migrate to region j if

$$U_j - c_{ij} = \max_k \{U_k - c_{ik}\}$$

[4] In the United States during the 16 years from fiscal year 1989 to 2004, 92% satisfy the transitivity condition for interstate migration with $n = 51$, and 99% satisfy the transitivity condition for interregional migration $n = 9$.

where c_{ij} is the migration costs from i to j which comprise non-pecuniary fixed costs. Assume further that individuals are heterogeneous in their perception of the attributes and characteristics associated with a particular region. Suppose the intertemporal utility is $U_i = u_i + \varepsilon_i$, where u_i is the deterministic intertemporal utility and ε_i is the random variable, which is independently and identically distributed across individuals with zero mean. We assume that the distribution of ε_i is double exponential so that the choice probability is given by the logit

$$P_{ij} = \frac{\exp\left(u_j - c_{ij}\right)/\alpha}{\sum_{k=1}^{n} \exp\left(u_k - c_{ik}\right)/\alpha}$$

where the randomness parameter α expresses the degree of heterogeneity in regional preference.

Figure 2 displays the density function of individual preference for two regions. The flatter density function is associated with the larger value of α, which implies heterogeneous preference and strong attachment to a particular region. On the other hand, the steeper density function with the smaller α implies greater homogeneous preference with greater sensitivity to the utility differentials than attachment to a region. An individual with positive ε_{ji} values region i relative to region j, i.e., the individual exhibits greater attachment to region i. Residents in i (resp. j) with $\varepsilon_{ji} > u_i - u_j + c_{ij}$ (resp. $\varepsilon_{ji} < u_i - u_j - c_{ij}$) migrate from i to j (resp. j to i): their share is represented by the right (resp. left) shaded area. Otherwise, they would stay in i (resp. j), and the share of these individuals is 1 minus the right (resp. left) shaded area. According to the first stylized fact on migration costs, the shaded areas must be small enough. The net migration is given by the difference between the two shaded areas, which must be smaller according to the second stylized fact on seemingly wasteful migration versus net migration.

It should be noted that P_{ij} is the probability when an individual in region i chooses region j if she receives *equal* opportunities with respect to, say, job offer or admission to a school. However, such opportunities are considered to be proportional to the size of the labor market in the destination region, L_j. Since the number of potential migrants in region i is L_i, the gross migration is specified as the product of these terms:

$$m_{ij} = L_i L_j P_{ij}. \tag{3}$$

Note that (3) is free from the aggregation of regions because gross migration is proportional to the sizes of both origin and destination.[5] Such gravity-type modeling is popular in dealing with interzonal traffic flows (Wilson, 1970) as well as interregional migration in the literature. In order to understand the aggregation problem, consider the case that the three regions are symmetrically located with an identical size and utility level. Then, it follows that $m_{ij} = m$ for

[5] Logit model of $m_{ij} = L_i P_{ij}$ does not satisfy the property.

all $i \neq j$. If regions 1 and 2 were to be aggregated, then gross migration from the aggregated region would be doubled. This is simply because the supply of migrants would be artificially doubled. Similarly, gross migration to the aggregated region would also be doubled because demand for migrants is artificially doubled.

Assume further that $c \equiv c_{ij} - c_{ii}$ is large enough based on the first stylized fact on migration costs. This means that the fixed costs of migration are large relative to the distance-related costs of migration. This is reflected by the fact that transport costs of migration are small relative to non-transport costs, such as adjustment costs in a destination region. Then, (3) is rewritten as

$$\begin{aligned} m_{ij} &= L_i L_j \frac{\exp\left(u_j - c_{ij}\right)/\alpha}{\exp\left(u_i - c_{ii}\right)/\alpha + \sum_{k \neq i} \exp\left(u_k - c_{ik}\right)/\alpha} \\ &\approx L_i L_j \frac{\exp u_j/\alpha}{\exp\left(u_i + c\right)/\alpha} \end{aligned} \tag{4}$$

Dividing m_{ij} by m_{ji}, and taking logarithm, we obtain

$$\log \frac{m_{ij}}{m_{ji}} = 2\left(u_j - u_i\right)/\alpha. \tag{5}$$

We adopt the regression equation (5), because it does not involve migration costs c_{ij}, and because it is consistent with the three stylized facts on interregional migration.[6]

It should be emphasized that involving migration costs c_{ij} raises a serious problem in estimating the utilities. In order to understand the problem, imagine a case that the three regions are located such that $(c_{12}, c_{23}, c_{13}) = (\bar{c}, \bar{c}, 2\bar{c})$ so that region 2 is the center and the other regions are peripheries. For simplicity, assume identical size $L_i = 1$, identical utility level $\bar{u}$, and $\alpha = 1$. Then, from (4),

$$\begin{aligned} m_{12} - m_{21} &= \frac{e^{u_2 - \bar{c}}}{e^{u_1} + e^{u_2 - \bar{c}} + e^{u_3 - 2\bar{c}}} - \frac{e^{u_1 - \bar{c}}}{e^{u_1 - \bar{c}} + e^{u_2} + e^{u_3 - \bar{c}}} \\ &= \frac{e^{\bar{c}} - 1}{\left(e^{\bar{c}} + 2\right)\left(e^{2\bar{c}} + e^{\bar{c}} + 1\right)} > 0 \end{aligned}$$

[6] There are other specification candidates, such as the first-order Taylor series expansion

$$m_{ij} - m_{ji} = s\left(u_j - u_i\right)$$

the replicator dynamic by Wall (2001)

$$\frac{m_{ij} - m_{ji}}{L_i L_j} = s\left(u_j - u_i\right)$$

and Douglas (1997)

$$\frac{m_{ij}}{L_i} - \frac{m_{ji}}{L_j} = s\left(u_j - u_i\right)$$

where s is a constant. However, none of them exhibited a better fit in terms of adjusted R^2 than (5), and the last one does not satisfy the third stylized fact of payoff monotonicity (1).

holds. This shows that in the presence of migration costs, *net migration always occurs even though there is no utility differential.* This means that payoff monotonicity does not hold insofar as migration costs are present.

Conversely, what if net migration is zero ($m_{ij} = m_{ji}$) and the utility of the center is different from that in the peripheries ($u_2 \neq u_1 = u_3$)? It can be readily shown that the utility of the center is *lower* than that in the peripheries under zero net migration, which is indeed problematic. Such a problem arises because utilities are not identifiable in the specification of (4): the utility in destination region j differs from where a migrant comes from. In order to avoid the identification problem, *it is imperative not to include distance-related migration costs* c_{ij} *in the utility estimation.*

4 Estimation method

Usually, in estimating (5), each utility u_i is expressed by a function of regional attributes, such as wages, prices, rent, amenities, and so on, which are referred to as determinants of migration. However, the choice of attributes and the functional form of the utility are ad hoc. In fact, according to Greenwood (1975, 1985), there are diverse determinants of interregional migration, all of which are not used as independent variables due to econometric problems, such as multicollinearity. Furthermore, there are a variety of specifications of (indirect) utility functions in the literature (Baldwin *et al.*, 2003). We therefore do not use a set of regional attributes as independent variables with a particular utility function. Instead, we directly estimate the values of the utilities themselves by using regional dummy variables only as independent variables. This can be performed by the following OLS regression:

$$\log \frac{m_{ij}}{m_{ji}} = \sum_{k=2}^{n} b_k D_k + e_{ij} \qquad i, j = 1, \ldots, n \text{ and } i \neq j \tag{6}$$

with no intercept for each year. The regression coefficient is $b_k = u_k/\alpha$, the residual term is e_{ij}, and the dummy variable is

$$D_k = \begin{cases} 1 & \text{if } k = j \\ -1 & \text{if } k = i \\ 0 & \text{otherwise.} \end{cases}$$

There are ${}_nC_2$ regional pairs, which is the number of observations. Note that D_1 should be dropped from the RHS of (6) because of linear dependency $\sum_{k=1}^{n} D_k = 0$. Since $b_k = b_k - b_1 = (u_k - u_1)/\alpha$ holds for $k = 2, \ldots, n$, we are able to estimate the utility differentials up to the multiple of the heterogeneity parameter α.

If a person migrates from regions i to j and migrates back from j to i in the same period, it is not counted as gross migration. Such return migration often takes place for a longer period, e.g. five years in the censuses of Japan and the United States. In this case, the number of return migrants should be subtracted from the numerator and denominator of the LHS of (6). However, this is not possible insofar as the number of return migrants is unknown. We therefore use the annual data instead of the five-year census data in estimating (6).

4.1 Interregional migration in Japan

Due to the residence register system of Japan, the interprefectural migration matrix has been published annually since 1954. Utilizing these annual data for 1954-2005, we aggregate the prefectures into $n = 9$ regions (see Appendix (i) for details of the aggregation). This is because migrants would not correctly recognize the distinction between regional utilities without regional aggregation, i.e., $n = 47$ prefectures, and because large metropolitan areas extend through several prefectures.

We ran regression (6) for $n = 9$ regions for each year from 1954 to 2005. For example, the result for the year 2005 is summarized as follows:

	b_2	b_3	b_4	b_5	b_6	b_7	b_8	b_9
coefficient	0.02	0.38	0.10	0.33	0.19	0.13	0.04	0.10

adjusted $R^2 = 0.90$, SD for each $b_i = 0.029$, NOB $= {}_nC_2 = 36$

We observe that the utility in region 3, which includes Tokyo, is the highest and that in region 5, which includes Nagoya, is the second highest. Furthermore, most of the coefficient estimates in different years are similar, the adjusted R^2 are high (the average is 0.79), and the common standard errors are small relative to the coefficient estimates. This confirms the adequacy of the specification (6).

Assuming that heterogeneity parameter α does not vary over time, we set $\alpha = 1$. In order to see the change in the regional ranking, define the standardized utility as

$$\hat{u}_i \equiv u_i - \bar{u} = b_i - \bar{b}$$

where $\bar{u} = \frac{1}{n}\sum_j u_j$ is the average utility and $\bar{b} = \frac{1}{n}\sum_j b_j = (\bar{u} - u_1)/\alpha$ is the average coefficient estimate. Figure 3 presents the values in the standardized utility $\hat{u}_i$ for region $i = 1, \ldots, 9$ for the study period. Although there are a few peaks and troughs, we can observe a regular tendency in which the standardized utilities roughly converge over time.

In order to confirm this conjecture, define the utility differential (UD) index as the standard deviation of the utility:

$$\text{UD} \equiv \sqrt{\frac{1}{n}\sum_i (u_i - \bar{u})^2} = \alpha\sqrt{\frac{1}{n}\sum_i \left(b_i - \bar{b}\right)^2}.$$

Figure 4 shows the decreasing trend of the UD with the simple average of 0.17. In fact, if we conduct a simple regression of the UD on the year, then the regression coefficient of the year is negative and highly significant with a t-value of -9.45. We thus conclude that *the interregional utility differential in Japan has been roughly decreasing after World War II.*[7]

Why is the adjustment speed toward interregional equilibrium ($u_i = \bar{u}$, $\forall i$) so slow? While the interregional price differentials in traded goods diminish quickly, the interregional utility differential converges at a relatively slower rate. This may be ascribed to the first stylized fact on migration costs: there exist many constraints in migration adjustment. For example, students are unable to enter interregional job markets until graduation, recruitment examinations do not take place very frequently, and the decision to migrate to a distant place takes time. Note that while interregional equilibrium of commodity prices, wages and rent is attained in the short run, spatial equilibrium of equalizing interregional utility levels would be attained in the long run as is discussed in new economic geography (Krugman, 1991).

Why do the interregional utility differentials diverge in some periods, such as prior to 1960, 1975-85, and after 1995, which exhibit positive slopes in Figure 4? While there is a market mechanism through migration adjustments to diminish the utility differentials, each utility itself changes continually due to rapid changes in regional economic environments, e.g. rapidly growing sectors in some regions. The changes are so rapid that the migration toward interregional equilibrium may not catch up. As Evans (1990) argues, "it is difficult to generate a convincing theoretical argument which would reconcile the assumption with continuing existence of persistent patterns of migration."

5 Robustness checks

In this section, we check the robustness of the model (5) in three different ways.

[7] As a robustness check, we also analyzed the case of $n = 46$ prefectures. It was revealed that the values of the UD is very similar: the correlation between the UD with $n = 9$ and the UD with $n = 46$ was 0.99. However, the average adjusted R^2 of the 52 years of the period 1954-2005 is 0.38 in the case of $n = 46$. This is possibly because some prefectures are so small that big metropolitan areas like Tokyo are divided into a few prefectures.

5.1 Interindustry movement in Japan

Our model would be applicable to any mobility between subgroups insofar as it satisfies the three stylized facts, which may be true for interindustry mobility. In order to examine this, we constructed Japanese interindustry mobility matrices for the years for which the Employment Status Survey was conducted, namely, 1968, 1971, 1977, 1982, 1987, 1992, 1997 and 2002 (see Appendix (ii) for the explanation of the industries). We ran the same regression (6) for $n = 7$ industries. The result for the year 2002 is as follows:

	b_2	b_3	b_4	b_5	b_6	b_7
coefficient	0.15	0.44	0.17	0.27	0.63	0.28

adjusted $R^2 = 0.65$, SD for each $b_i = 0.106$, NOB $= {}_nC_2 = 21$

In the remaining seven years, the coefficient estimates are more or less similar, and the adjusted R^2 are high (the average is 0.65), implying the appropriateness of our model. The value of the UD with $\alpha = 1$ is between 0.19 and 0.33 with a simple average of 0.26. Unlike the interregional utility differential, the interindustry utility differential does not exhibit convergence during the study period.

5.2 Interregional migration in the United States and Canada

The second check is conducted by using the data of the interregional migration flows in the United States and Canada, where the stylized facts seem to hold as well. In the United States, 51 states are aggregated into 9 regions: 1 New England, 2 Middle Atlantic, 3 East North Central, 4 West North Central, 5 South Atlantic, 6 East South Central, 7 West South Central, 8 Mountain, and 9 Pacific (see Appendix (iii) for these definitions). Then, we have 9×9 interregional migration matrices for 16 years from the fiscal year 1989 to 2004. As in the Japanese case, we ran regression (6) each year from 1989 to 2004 for the United States. For example, the result for 2004 is as follows:

	b_2	b_3	b_4	b_5	b_6	b_7	b_8	b_9
coefficient	−0.01	0.01	0.14	0.50	0.43	0.33	0.49	0.18

adjusted $R^2 = 0.87$, SD for each $b_i = 0.053$, NOB $= {}_nC_2 = 36$

In the other years, the coefficient estimates are similar, the average adjusted R^2 is 0.90, and the common standard error is small relative to the coefficient estimates,all of which verify the validity of our model. Next, the value of the UD with $\alpha = 1$ is shown to range from 0.14 to

0.27 with a simple average of 0.20. Regressing the UD on the year yields a negative regression coefficient with a t-value of -2.94, thus implying the converging trend in the interregional utility differential.[8]

We also run the regression using the interprovincial migration data of Census Canada with $n = 13$ provinces during the five year period of 1996-2001 (Appendix (iv)), we do the same. The result is as follows:

	b_2	b_3	b_4	b_5	b_6	b_7	b_8	b_9	b_{10}	b_{11}	b_{12}	b_{13}
coefficient	1.02	0.90	0.90	0.57	1.08	0.80	0.86	1.65	0.94	0.54	0.65	0.70

adjusted $R^2 = 0.72$, SD for each $b_i = 0.124$, NOB $= {}_nC_2 = 78$

If the provinces are aggregated into $n = 5$ regions according to the definitions by Rogers, Raymer and Newbold (2003), we obtain an adjusted $R^2 = 0.78$. Setting $\alpha = 1$, the computed UD is 0.38 for $n = 13$, and 0.34 for $n = 5$. These results for Canada are rather similar to those for Japan and the United States. Thus, the robustness of our model is also verified by the American and Canadian data.

5.3 Comparison with real income

The final robustness check involves the comparison of the utility estimates with the real income in order to examine whether the estimated utilities are reliable or not. If the utilities are correctly estimated by (6), then they would be highly correlated with the per capita real income. In order to verify this, we run the pooled regression:

$$u_{it} = \beta I_{it} + \sum_{j=2}^{n} \gamma_j R_j + \sum_{t=t_1}^{t_2} \delta_t Y_t + e_{it} \qquad i = 1, \ldots, n \text{ and } t = t_1, \ldots, t_2$$

without intercept, where I_{it} is the per capita real income defined by the per capita nominal income divided by the consumer price index in region i and year t, R_j is the regional dummy, Y_t is the year dummy, and β, γ_j and δ_t are the regression coefficients. The regional dummies are needed because there are non-economic factors, such as regional amenities, which are not captured by the per capita real income. The year dummies are also needed because booms and recessions are not controlled by the per capita real income.

We first ran the pooled regression using Japanese data. Due to the constraint of the consumer price index, the data was available for the period of 1979-2002 (Appendix (v-vi)). The pooled regression produces a good fit. Almost all the coefficients are significant, and the adjusted R^2 is

[8] We also analyzed the case of $n = 51$ states. It was revealed that the values of the UD are very similar: the correlation between the UD with $n = 9$ and the UD with $n = 51$ is 0.95. However, the average adjusted R^2 of the 16 years of 1989-2004 is 0.65 in the case of $n = 51$. Again, this is possibly because some states are too small.

0.84. Since the data on the per capita income is available for the longer period of 1955-2002, we also ran the same regression by replacing I_{it} with the per capita nominal income for the study period. This regression performed well too: almost all the coefficients are significant, and the adjusted R^2 is 0.75.

We then ran the pooled regression using American data by replacing I_{it} with the per capita nominal income for the study period of 1989-2004 (Appendix (vii)). Again, this regression performed very well: almost all the coefficients are significant, and the adjusted R^2 is 0.90.

Hence, we conclude that these results support the validity of the utility estimation.

6 Conclusion

Three stylized facts in migration behavior are revealed from the Japanese data from 1954 to 2005. Based on the stylized facts, we developed an operational model for analyzing interregional migration as well as intersectoral mobility. Unlike the previous literature, our model does not need to impose any ad hoc assumptions on the utility function. Using the interprefectural migration data in Japan, we estimated the interregional utility differentials. Estimation results suggest that our model produces an extremely good fit for each year, and that interregional utility differentials have been converging particularly until the late 1970s.

In order to check the robustness of our model, we first applied the model to the intersectoral mobility in Japan and estimated the intersectoral utility differentials, and confirmed the good fit of our model. We then ran regressions using interregional migration flows in the United States and Canada, confirming its validity in this context as well. Finally, we regressed the utility estimates on the real wages, and obtained high correlations. Hence, the model presented in this paper is considered to be useful in estimating utility differentials between regions as well as between industry sectors.

Data Appendix

(i) **Interprefectural migration in Japan** is in the Report on Internal Migration in Japan (Japanese Statistics Bureau) for 1954-2005. Following Ishikawa (2001), 47 prefectures are aggregated into 9 regions: region 1 is Hokkaido; region 2 consists of Aomori, Iwate, Miyagi, Akita, Yamagata, Fukushima and Niigata prefectures; region 3 consists of Ibaraki, Tochigi, Gumma, Saitama, Chiba, Tokyo, Kanagawa, Yamanashi and Nagano prefectures; region 4 consists of Toyama, Ishikawa and Fukui prefectures; region 5 consists of Gifu, Shizuoka, Aichi and Mie prefectures; region 6 consists of Shiga, Kyoto, Osaka, Hyogo, Nara and Wakayama prefectures;

region 7 consists of Tottori, Shimane, Okayama, Hiroshima and Yamaguchi prefectures; region 8 consists of Tokushima, Kagawa, Ehime and Kochi prefectures; and region 9 consists of Fukuoka, Saga, Nagasaki, Kumamoto, Oita, Miyazaki and Kagoshima prefectures. Okinawa prefecture is excluded in the regressions.

(ii) **Interindustry mobility in Japan** is in the Employment Status Survey in Japan (Japanese Statistics Bureau, 1968, 1971, 1977, 1982, 1987, 1992, 1997 and 2002). Unlike the interregional migration, this is the sample survey of 440,000 households. 13 industries are aggregated into 7 industries: 1 manufacturing and mining, 2 construction, 3 electricity, gas, water, transportation and communication, 4 wholesale and retail trade, 5 finance, insurance and real estate, 6 services, and 7 government. Agriculture, forestry, hunting, fisheries, aquaculture and unclassified industry are excluded.

(iii) **Interstate migration in the United States** is in the State-to-State Migration Data (Internal Revenue Service, U.S. Department of the Treasury) for the fiscal year 1989-2004. 51 states are aggregated into 9 regions: region 1 is New England consisting of Connecticut, Maine, Massachusetts, New Hampshire, Rhode Island, and Vermont; region 2 is Middle Atlantic consisting of New Jersey, New York, and Pennsylvania; region 3 is East North Central consisting of East North Central, Illinois, Indiana, Michigan, Ohio, and Wisconsin; region 4 is West North Central consisting of Iowa, Kansas, Minnesota, Missouri, Nebraska, North Dakota, and South Dakota; region 5 is South Atlantic consisting of Delaware, District of Columbia, Florida, Georgia, Maryland, North Carolina, South Carolina, Virginia, and West Virginia; region 6 is East South Central consisting of Alabama, Kentucky, Mississippi, and Tennessee; region 7 is West South Central consisting of Arkansas, Louisiana, Oklahoma, and Texas; region 8 is Mountain consisting of Arizona, Colorado, Idaho, Montana, Nevada, New Mexico, Utah, and Wyoming; and region 9 is Pacific consisting of Alaska, California, Hawaii, Oregon, and Washington.

(iv) **Interprovincial migration in the Canada** is in Profile of the Canadian Population by Mobility Status: Canada, a Nation on the Move (2001 Census: Analysis Series, Catalogue no. 96F0030XIE2001006) for the period of 1996-2001.

(v) **Consumer price index in Japan** is in the Annual Report on the Consumer Price Index, Statistics Bureau, Ministry of Internal Affairs and Communications for 1979-2002.

(vi) **Per capita income in Japan** is the prefectural income divided by prefectural population for 1955-2002. The data are in the Annual Report on Prefectural Accounts, Economic and Social Research Institute, Cabinet Office.

(vii) **Per capita income in the United States** is the state personal income divided by state population for 1989-2004. The data are in the Regional Economic Accounts, Bureau of

Economic Analysis, U.S. Department of Commerce.

References

[1] Anderson, S.P., A. de Palma and J.-F. Thisse (1992) *Discrete Choice Theory of Product Differentiation*, Cambridge (Mass.): MIT Press.

[2] Baldwin, R., R. Forslid, P. Martin, G. Ottaviano and F. Robert-Nocoud (2003) *Economic Geography and Public Policy*, Princeton: Princeton University Press.

[3] Cragg, M. and M.E. Kahn (1997) New estimates of climate demand: evidence from location choice, *Journal of Urban Economics* 42, 261-284.

[4] Crozet, M. (2004) Do migrants follow market potentials?, *Journal of Economic Geography* 4, 439-458.

[5] Cushing, B. and J. Poot (2004) Crossing boundaries and borders: Regional science advances in migration modelling, *Papers in Regional Science* 83, 317–338.

[6] Douglas, S. (1997) Estimating relative standards of living in the United States using cross-migration data, *Journal of Regional Science* 37, 411-436.

[7] Evans, A.W. (1990) The assumption of equilibrium in the analysis of migration and interregional differences: A review of some recent research, *Journal of Regional Science* 30, 515-31.

[8] Greenwood, M.J. (1975) Research on internal migration in the United States: a survey, *Journal of Economic Literature* 13, 397-433.

[9] Greenwood, M.J. (1985) Human migration – Theory, models, and empirical studies, *Journal of Regional Science* 25, 521-544.

[10] Grubel, H.G. and P.J. Lloyd (1971) The empirical measurement of intra-industry trade, *Economic Record* 47, 494-517.

[11] Ishikawa, Y. (2001) *Studies in the Migration Turnarounds*, Kyoto: Kyoto University Press.

[12] Kahn, M.E. (1995) A revealed preference approach to ranking city quality of life, *Journal of Urban Economics* 38, 221-235.

[13] Krugman, P. (1991) Increasing returns and economic geography, *Journal of Political Economy* 99, 483-499.

[14] Pons, J., E. Paluzie, J. Silvestre and D.A. Tirado (2007) Testing the new economic geography: Migrations and industrial agglomerations in Spain, *Journal of Regional Science* 47, 289-313.

[15] Rogers, A., J. Raymer and K.B. Newbold (2003) Reconciling and translating migration data collected over time intervals of differing widths, *Annals of Regional Science* 37, 581–601.

[16] Samuelson, P.A. (1948) Consumption theory in terms of revealed preference, *Economica* 15, 243-253.

[17] Tabuchi, T. and J.-F. Thisse (2002) Taste heterogeneity, labor mobility and economic geography, *Journal of Development Economics* 69, 155-177.

[18] Tiebout, C.M. (1956) The pure theory of local expenditures, *Journal of Political Economy* 64, 416-424.

[19] Wall, H.J. (2001) Voting with your feet in the United Kingdom: Using cross-migration rates to estimate relative living standards, *Papers in Regional Science* 80, 1-23.

[20] Weibull, J.W. (1995) *Evolutionary Game Theory*, Cambridge (Mass.): MIT Press.

[21] Wilson, A.G. (1970) *Entropy in Urban and Regional Modelling*, London: Pion.

Figure 1: Interprefectural gross and net migration flows

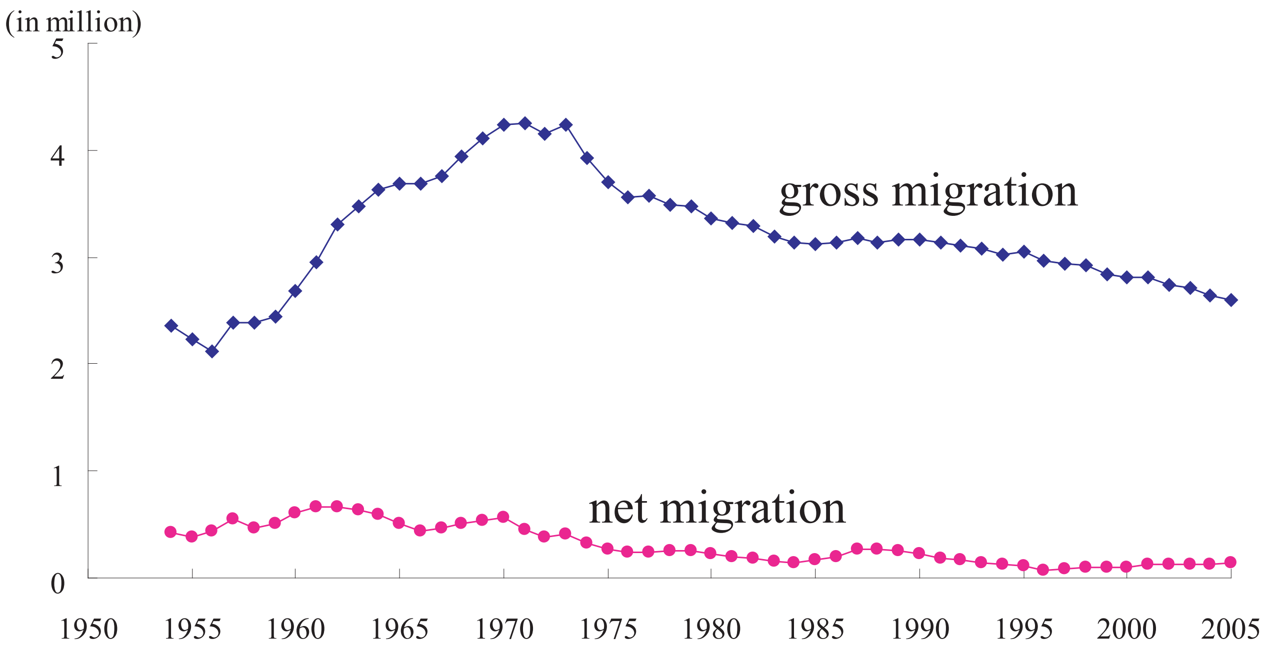

Figure 2: Probability density when $u_j > u_i$

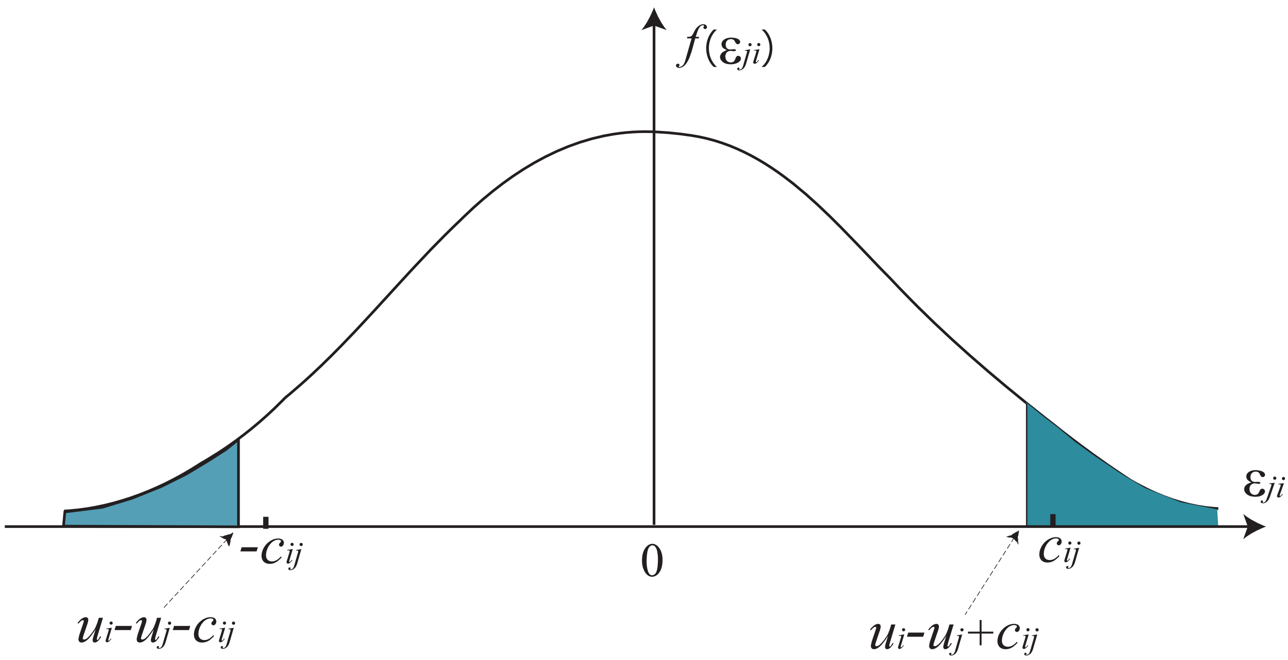

Figure 3: Estimated utility differentials in 9 regions

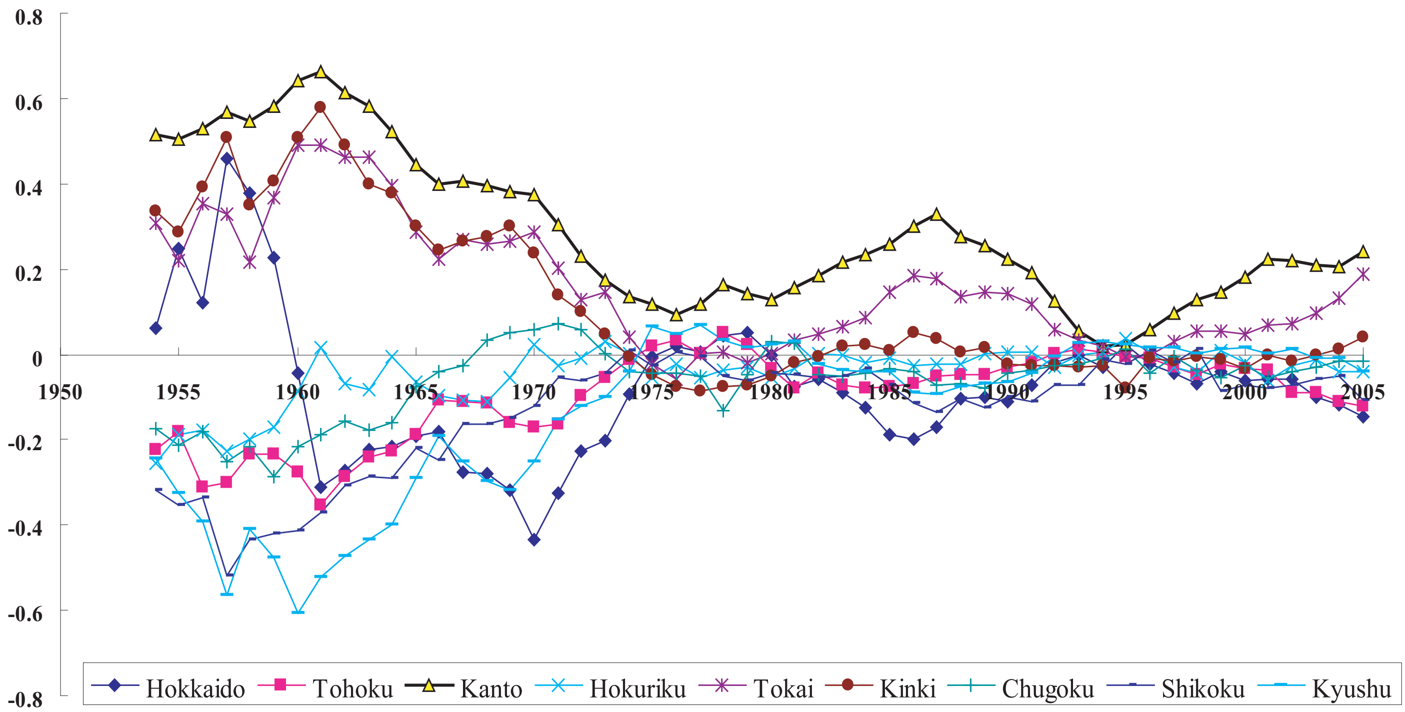

Figure 4: Utility differential index