Wages Differentials and Interregional Migration in the U.S.: an Empirical Test of the "Option Value of Waiting" Theory*

Roberto Basile[1], Jaewon Lim[2]

[1] Institute for Studies and Economic Analyses (ISAE) and University of Macerata, Italy
Piazza dell'Indipendenza, 4 00185–Rome, Italy. (e-mail: r.basile@isae.it)
[2] Regional Economics Applications Laboratory (REAL), University of Illinois at Urbana-Champaign, USA
607 S Mathews Ave, Urbana, IL 61801 USA. (e-mail: jlim2@uiuc.edu), Corresponding author

Abstract: The '*option value of waiting*' theory applied to interregional migration predicts that a potential migrant actually moves only when the wage differential between origin and destination places exceeds a certain threshold, which might be much higher than the Marshallian trigger. In this paper we exploit the panel structure of a dataset on interregional migration among nineteen MSAs in the US from 1993 to 2001 to estimate a modified dynamic gravity model of migration. In particular, using both semi-parametric and GMM estimators (taking into account possible endogeneity of the explanatory variables), we find robust evidence of a non-linear relation between migration and wage differentials. With a wage differential smaller than a certain threshold, people rarely move controlling for the other socioeconomic variables. Only beyond the threshold, the interregional migration grows rapidly proving an important role of the option value of waiting in migration decision process.

Keywords: wage differentials, interregional migration, option value of waiting, modified gravity model, semi-parametric analysis, GMM

JEL classification: C14, C33, J30, R23

---

* The present paper has been written while Roberto Basile was visiting the Regional Economics Applications Laboratory at UIUC. A particular thank is due to Geoffrey J.D. Hewings both for providing hospitality during that period and for all the comments and suggestions helpful to improve the paper from both of the authors.

1. Introduction

The empirical studies on migration are enormous and growing, with most studies focusing on inter-regional migration or, in other words, place-to-place migration (for a review, see Greenwood, 1997). Generally speaking, these studies use different modifications of the Harris-Todaro (1970) model[1] to explain either net migration or gross migration.[2] In most cases, it would be more desirable to model gross in- and out-migration than to model net migration due to the volatility of net migration and, more importantly, the fallacy of net migration rates (Plane and Rogerson, 1994). Rogers (1990) demonstrated the violation of the demographer's principle when using the net migration rates, since the at-risk population for net migration which is used as a denominator is composed of population in specific sets of origin, and destination places included in the study. However, the relevant at-risk population for in-migration is not the population from specific sets of origin but all those who are not in the destination under consideration (Plane and Rogerson, 1994). Using a modified gravity model with gross place-to-place migration flows as dependent variable overcomes this problem. Starting from mid 1980s, dynamic specifications of the gravity migration model have been in common due to the availability of longitudinal dataset on migration (Molho, 1984). Interregional migration can be partly explained by both temporal persistence and temporal volatility. The extent of which factor play a greater role depends on socioeconomic characteristics of origin and destination places, personal attributes of the two places and possible temporal variations in national and regional economic conditions.

[1] Such as the models proposed by Pissarides and McMaster (1990), Decressin (1994) and Oswald (1990).
[2] Net migration is defined as total inflows in a region less total outflows from a region, while gross migration is defined as the number of immigrants in the region of destination.

Surprisingly enough, most of migration studies using aggregate data and based on (static or dynamic) modified gravity models have implicitly assumed that all regions obey a common log-linear specification. In particular, researchers in interregional migration generally assume a log-linear relationship between the volume of migration and the differences in regional economic conditions of origin and destination places. However, this assumption does not always hold, especially when an "option value of waiting" influences the migration decision process. As Burda (1993) and Parikh and Van Leuvensteijn (2002) pointed out, the migration decision can be sensitive to the option value of waiting, since it is characterized by the following features: a) a fixed sunk cost, b) uninsurable uncertainty and c) the possibility of waiting and postponing the decision and therefore, postponing the payment of the fixed costs. Due to the possible presence of option value of waiting, a vulnerable migrant chooses to actually move only beyond some thresholds in terms of the wage differences between the region of destination and that of origin, rather than just moving when this difference is positive. Thus, a non-linear relationship between migration and wage differentials may be the most plausible outcome.

This paper studies the dynamic patterns of interregional migration among the Metropolitan areas (MSAs) in the United States during 1990s. The nineteen MSAs included in this study are composed of the ten largest MSAs based on the population from Census 2000, the largest seven MSAs in Midwest, and two MSAs in California, ranked below the top ten.[3] The interregional migration patterns are formulated in an economic-demographic dynamic gravity model. A new dataset is exploited to empirically investigate the relationship between interregional migration flows and wages in the USA.[4] In particular, the dynamic gravity model is estimated using annual

[3] The list of 19 MSAs included in this study can be found from the appendix with their locations displayed on a map.
[4] Most of the studies on interregional migration in the USA have used Census data from the Current Population Survey (CPS). These data do not have a sequential time dimension, so that they do not allow testing a dynamic

interregional migration data from Internal Revenue Service (IRS) among the nineteen selected Metropolitan Statistical Areas (MSA) in the United States for the time period between 1992 and 2001.

As a starting point, the possible presence of non-linearity in the relationship between migration flows and wage differentials is investigated (i.e. the option value of waiting hypothesis is tested) by applying a non-parametric, panel, fixed effect regression model. Then, the information from the non-parametric analysis is used to properly specify the functional form of a parametric dynamic panel data model and the GMM-System methodology (Blundell and Bond, 1998) is applied in order to control for all potential endogeneity sources: simultaneity and measurement errors problems.

In the following section, a review of the literature on migration decision is provided. Section 3 describes the option value of waiting theory applied to the migration decision process. Section 4 specifies the empirical model to test the ‘option value of waiting' hypothesis and presents the empirical evidence. Section 5 reports some conclusions.

2. Migration Decision: A Literature Review

In this section of paper, the literature on the determinants of migration decision-making is reviewed, mainly focusing on the roles of wages and unemployment, t amenities, and the role of time, especially the issues related to the temporal aspects in migration decision. In the last part of this section, the concept of ‘option value of waiting,’ suggested by Dixit (1992) and Pindyck

---

relationship. Instead, we use an annual panel dataset which makes it possible to exploit both the time-series and cross-region variation in immigration inflows.

(1991) in the field of investment decision, will be reviewed as a framework for the migration decision following Burda (1993).

2.1 The role of wage and unemployment differentials

Hicks (1932) and Makower *et al.* (1938, 1939, 1940) attributed the economic incentives, expressed in the regional differences between origin and destination regions, to the main causes of interregional migration. On the one hand, Hicks saw the demand and supply of labor as being mediated by fluctuations in wages in the classical tradition and stated that "*... differences in net economic advantage, chiefly differences in wages, are the main cause of migration*" (Hicks, 1932; p.76). On the other hand, Makower *et al.* in their series of papers (1938, 1939, 1940) emphasized on the roles of unemployment differentials and distance. Explaining "*relative unemployment discrepancies*", the authors did not specifically formulate a gravity law of spatial interaction. However, they described similar concepts by stating "*Quite a close relationship was found between discrepancies in unemployment rates and migration of labor where allowance was made for the size of the insured population and the distance over which migrants had to travel*" (Mankower *et al.*, 1938: p. 118).

Sjaastad (1962) modeled migration as an investment process in human capital (Sjaastad, 1962) and this had been the foundation of the dominant economic theory of migration. Hart (1975) outlined the human capital approach as follows: potential migrants evaluate the expected utility, E (U) less the expected discounted costs (moving costs), C for each of the possible destinations $j = 1, \ldots, N$ (including the original location i) and select to live in the area with the highest net outcome, or net present value. The expected utility over time period, T, can be written as equation (1) shown below:

$$E\{U[R_{ij}(0)]\} = \int_0^T \exp(-rt)U[R_{ij}(t)]dt = \frac{R_{ij}}{r} \tag{1}$$

where,

R_{ij}: wage flow differentials between origin and destination places ($logW^D$-$logW^O$) per unit of time (say per year)

T: time horizon (up to retirement) over which the individual calculates returns from wage flows

r: subjective discount rate, which depends on the age structure of the population at risk and the expected discounted costs can be formulated as equation (2) below:

$$E[C_{ij}(0)] = \int_0^T \exp(-rt)[C_{ij}(t)]dt = f \tag{2}$$

The costs include both monetary, mainly direct and moving costs, and psychic costs including the loss of attachment to relatives and friends in origin places. The human capital model clearly stated migration as a utility maximizing process in the face of economic opportunity differentials which represent potential for utility gains. Through migration process, the existing economic opportunity differentials will diminish with the perfect information in a completely efficient labor market. As mentioned earlier, Hicks (1932) already shed light on these economic opportunity differences. In this sense, a disequilibrium perspective is evident in his model. For the next 20 years after Sjasstad first introduced the human capital model in 1962, this model provided a foundation and a framework for economists to discover the determinants of migration.

The net present value (NPV) is the difference between the expected utility and the expected discounted cost of moving:

$$NPV = \frac{R_{ij}}{r} - f \tag{3}$$

In figure 1, the straight line, NPV, represents the net present value of the migration decision as defined in equation (3) for each level of current wage differential given on the horizontal axis.

Under conditions of 'certainty' (about the wage differential flows), migration occurs 'immediately' when $\frac{R_{ij}}{r} > f$. An increase in the fixed cost, f, raises the Marshallian triggers $\overline{R}_{ij} = rf$.

<<Insert Figure 1 here>>

Later in late 1960s and 1970s, an economic model of interregional migration was laid out by Harris and Todaro considering both wage differences and probabilities to find a job, expressed in differentials in unemployment rates between origin and destination places (see Todaro, 1969; Harris and Todaro, 1970 and Todaro: 1976). In particular, the Harris and Todaro (1970) model is considered as a starting point for the modern analysis of interregional migration. In this model, risk neutral individuals with complete information take a decision to move on the base of a net present value calculation. More specifically, the decision to migrate depends on the expected income calculated on the base of a cost-benefit analysis which includes the probability to find a job in the destination. Originally, this model was oriented to explain the phenomenon of '*over-crowding*' and increasing unemployment in urban areas of the less developed countries, that is the movement of a large share of the work force (mainly young people) from rural low-wage areas towards urban and industrialized high-wage areas.

In the Harris and Todaro model, nominal wages in the urban industrial sectors are not completely flexible, rather they are rigid on the downside. The existence of a minimum wage influences the expectations on income of out-migrants. However, the existence of wage rigidity also generates unemployment in the urban areas. In fact, a worker may experience a period of unemployment or underemployment before he or she starts to earn the urban wage. Rational workers take into account this possibility in the calculation of their permanent income. Therefore, younger workers are more likely to migrate from rural to urban areas, since they have a life horizon long enough to discount the waiting time during which they might be unemployed or underemployed. The expected (or permanent) income of workers, and thus their incentive to migrate from rural areas to urban areas, is therefore an inverse function of the rural areas (or origin places) population age.

In the Harris-Todaro model, individuals calculate the expected income conditional on the probability to find a job that can be approximated by the unemployment rate in the destination place, u_j. Thus, the net present value (NPV) can be expressed as follows:

$$NPV = \frac{R_{ij}(1-u_j)}{r} - f \tag{4}$$

Again, under conditions of 'certainty', migration occurs when $\frac{R_{ij}(1-u_j)}{r} > f$.

More recently, micro-based models have been developed with predictions partly in line with and partly different from those postulated by Harris and Todaro (1970). In these models, individuals or households maximize their expected utility function, comparing the gross benefit to migrate with the cost of leaving the origin places. Pissarides and McMaster (1990) have proposed, for example, a modified Harris-Todaro framework to explain net migration rates. In this framework, households calculate the gross benefit of remaining in the origin places and compare it with the gross benefit of migration. Migration occurs if the gross benefit of moving exceeds the cost to move. This cost is affected by the observed and unobserved characteristics of households randomly distributed within the population. The gross benefit to migrate depends, on the other hand, on variety of other factors: wage differentials, unemployment rate differentials and the characteristics of households (age and skill levels). If the wage level in a specific region increases more than elsewhere, the gross benefit to migrate into this region increases, while the gross benefit from out-migration from the region decreases. The net migration rate of that region therefore increases. The unemployment rate also influences migration flows. Unemployed workers have indeed higher mobility since they have less to give up compared to employed workers (even if they may have fewer assets to afford a move). If unemployment rate in a region increases, then the net migration rate of the region would decrease.

As mentioned earlier, in Harris-Todaro model the wage level and the unemployment rate tend to be combined in a single variable, $R_{ij}(1-u_j)$. Unlike, the wage level and the unemployment rate may enter the model specification separately as suggested by Pissarides and McMaster (1990). Thus, the net present value (NPV) can now be expressed as follows:

$$NPV = E\left\{U\left[R_{ij}, u_{ij}\right]\right\} - f \tag{5}$$

where,
u_{ij} is the unemployment rate differential between the origin and destination places.

2.2 The role of amenities

Other authors (for example Decressin, 1994, and Oswald, 1990) introduce "*amenities*" into the utility function of households, generally approximated by climate conditions, the availability of houses, hospitals and other public infrastructure that may influence the quality of life. Most of the research in migration has failed to estimate a model with appropriate amenity variables. Furthermore, due to the compensating effects of amenities for regional differentials, there may implicitly be endogeneity problems associated with wages or income. However, this problem has not been addressed very often (Greenwood, 1997).

Some earlier empirical studies explained the interaction between regional amenities and regional economic conditions, more precisely, income levels and unemployment rates. Graves (1979) illustrates that climatological amenity variables play important roles in the estimation of age- and race-specific net migration in the 1960s considering income and unemployment of places. Moreover, his results indicate that when the amenity variables are excluded, income is usually insignificant. On the other hand, when the amenity variables are included, income variables are more likely to have the statistically significant expected signs. However, subsequent studies have found a less important role for amenity related variables (Greenwood and Hunt, 1989).

However, location-specific amenities still play an important role in estimating migration. If desirable places with better amenities attract more firms due to the lower wage, employment will expand very rapidly in those areas. It is clear that an increased number of jobs attracts migrants and, to some extent, jobs will be created due to amenities. In this perspective, amenities still play an important role to attract migrants in indirect ways.

Considering relative amenity levels between origin and destination places, a_{ij}, the net present value (NPV) can now be expressed as follows:

$$NPV = E\left\{U\left[R_{ij}, u_{ij}, a_{ij}\right]\right\} - f \tag{6}$$

2.3 The effect of time

The importance of temporal aspects in migration analysis is reflected in the following issues: first, the length of response lags to market signals; secondly, life cycle effects and ‘state dependence’ of individuals; thirdly, the cyclical perspective of the overall volume of migration; fourthly, the persistence in migration behavior.

First, the response lags have several possible causes, such as those outlined by Molho (1986). There are delays in the diffusion of information, expectations of future benefits streams depending on weighted average of past trends, and significant adjustment lags between the decision to migrate and the actual move. Molho (1984) in his earlier paper emphasized the need to specify the regional push and pull factors in some general distributed lag formulation.

Secondly, there exist some temporal issues related to life cycle effects and 'state dependence'. According to Molho (1986), 'state dependence' refers to "*the situation where individuals' migration decisions are explicitly affected by previous locational decisions in their life history*". Moreover, migration propensities vary over the life cycle depending on some personal characteristics. For example, Rogers *et al.* (1978) described how the migration rates vary by age. Later, Rogers and Castro (1986) showed how a model schedule can represent the various migration rates during the labor force years, and during the pre- and post-labor force years.

Thirdly, the cyclical nature of migration can be explained by the temporal variation of socioeconomic conditions, such as wage differentials and unemployment differentials, either at national or local levels. For cross-section gravity models, the implicit constant measures the aggregate volume of migration for a given period, whereas regional push and pull factors measure the deviations in in- and out-migration for each area (Molho, 1984). In the case of dynamic models, more precisely, panel data analysis on migration, variation of the general national economic climate over time cause the implicit constant vary over time since the aggregate volume of migration varies with the business cycle. More importantly, the extent of the impact from temporal variations in national business cycles will vary by region based on the regional economic structures. Empirical studies found that the volume of migration is likely to vary counter-cyclically (Hart 1975; Gordon, 1985). Since liquidity constraints restrict human capital investment behavior, such as migration, in a recession period uncertain prospects are discounted more heavily. Also, employment opportunities are likelier to be less available during a recession.

Fourthly, according to the ‘network approach’ (Ghatak, et al., 1996; Bauer and Zimmermann, 1995), people that out-migrated in the past influence the choice of workers contemplating out-migration today, by reducing fixed costs and risks of entry and rendering the migration process easier to realize. This argument may suggest using a dynamic specification of the gravity model in order to reflect some degree of persistence.

Finally, one has to consider the role of uncertainty related to the temporally varying socioeconomic characteristics. Potential migrants do not have perfect insights on the future levels of wage and unemployment differentials and, thus, they are not always able to maximize their utilities from migration. As will be clarified in the next section, under these conditions, potential migrants have the option to postpone their decision to move: waiting for a certain amount of time enables them to reduce the risks connected to the presence of uncertainty. Consequently, the traditional decision criteria on migration - “whether to move” and “where to move” - may be expanded to include another decision criteria, “when to move” with the possible presence of ‘option value of waiting’ in interregional migration.

3 The ‘option value of waiting’ and the migration decision

Burda (1993) and Parikh and Van Leuvensteijn (2002) have recently suggested that the responses of migrants to wage differentials may be characterized by some non-linearities due to what they call an "option value of waiting".[5] In fact, migration behavior is characterized by the following features. First, migration entails fixed sunk costs that cannot be recouped if the action is reversed at a later time. Secondly, the economic environment is characterized by uncertainty, and information arrives gradually. Thirdly, there exists the possibility of waiting and postponing the decision to migrate; therefore, the decision on migration is composed of two parts, whether to move and when to move. Given these three features, waiting has some positive value since it reduces risks over time. Indeed, waiting for a certain amount of time enables a migrant not only to avoid the downside risk in wages over that interval, but also to realize the potential increases in wage differential.[6] In such an environment, migration occurs only when the wage differential exceeds the 'Marshallian trigger'[7] by a positive margin. In other words, due to the 'option to wait,' a potential migrant chooses to actually move only beyond some thresholds in terms of the wage differential between the region of destination and that of the origin, rather than just moving when the utility of moving (net of the fixed cost) is positive. Thus, a non-linear relationship between migration and wage differentials may be the most plausible outcome, as shown in Figure 2.

<<Insert Figure 2 here>>

In the previous section, we postulated that, under conditions of 'certainty' (about the wage differential flows), migration occurs 'immediately' when $\frac{R_{ij}}{r} > f$. Under condition of uncertainty, however, this is not true anymore. Suppose that the future wage differential flows

[5] The 'value of waiting' analysis was initially applied to the valuation and optimal exercise of financial options (see Dixit and Pindyck, 1994).

[6] It is important to note that the value of waiting is not related to risk-aversion, since it has been established under the assumption of risk neutrality; and in this sense the option value of waiting may be even consistent with the Harris-Todaro framework.

[7] The Marshallian trigger is the level of wage differential at which the expected net present value is zero.

are only imperfectly predictable from the current observation of R_{ij}. The probability distribution of future wage differentials is determined by the present, but the actual path remains uncertain. The probability law of evolution of R_{ij} can take many forms. We only suppose that in each period, R_{ij} can either increase or decrease by a fixed percentage. Suppose the Marshallian trigger is the starting point to consider migration. It would be still profitable to wait for a certain period of time for two reasons: first, in the case of increasing wage differentials, a migrant would be able to realize the potential wage increases in the future; secondly, in the case of decreasing wage differentials, a migrant would avoid the losses of wages by waiting. Consequently, even if $R_{ij} > \overline{R}_{ij}$ (the Marshallian trigger), the individual has the option to wait in order to avoid the mistake of moving and losing income and to realize possible future increases. Thus, waiting is valuable for a potential migrant.

On the other hand, the cost of waiting is the foregone wage income over the period of waiting. Thus, when the current wage differential, R_{ij}, is sufficiently higher than the Marshallian trigger, $\overline{R}_{ij}$, it would be unprofitable to wait any longer. In figure 2, point *E* represents a critical level of wage differential (higher than the Marshallian level) beyond which migration is always optimal. At this point, the net present value function for a potential migrant is tangent to the value of waiting function $V = V\left(R_{ij}\right)$, such that the migrant will be indifferent between the decision of waiting to migrate and the decision to migrate now. At all points beyond *E*, the decision to migrate will dominate the decision to wait as the former has higher utility than the latter.

In a nutshell, according to the 'option value of waiting' theory, a potential migrant actually moves only when the wage differential between origin and destination places exceeds a certain threshold point *E*. If this theory holds, a non-linear relation between migration and wage differentials should be observed from a cross-section or a panel data analysis. Put differently, we expect that for low levels of wage differentials, a worker would not move from the origin place, while after a threshold he or she would actually move. So, the most plausible outcome may be a positive relationship between wage differentials and migration only after a threshold (E) in R_{ij}.

4. Empirical evidence on the interregional migration in U.S.

In this paper, we exploit the panel structure of the IRS (Internal Revenue Service) dataset that provides information on the interregional population migration between 19 US MSAs for the period 1994-2001, by estimating a modified gravity model. This is an important extension of earlier studies of determinants of population migration that have focused on the cross-section variation within a single period using CPS data. We first perform a semi-parametric analysis of the dynamic gravity model to test the existence of a non-linear relationship between population migration and wage differentials (Section 4.1). Then, the information from this analysis is used to identify the polynomial (parametric) transformation of the dynamic panel migration model. Finally, the GMM methodology (Arellano and Bond, 1991) is applied in order to control for endogeneity (Section 4.2). Implications of the estimates are further discussed in Section 5.

4.1 A semi-parametric gravity model of interregional migration

As mentioned above, in order to identify the presence of non-linearities in the relationship between wage differentials and migration, we use the semi-parametric methodology. In particular, by using a particular version of the semi-parametric model that allows for additive components, we are able to obtain graphical representation of the relationship between wages and migration. Indeed, additivity ensures that the effects of each of the model predictors can be interpreted net of the effects of the other predictors, just as in linear multiple regression. The semi-parametric model can be written as

$$
\begin{aligned}
y_{ijt} &= \ln\left[\frac{M_{ijt}}{L_{it}L_{jt}}\right] \\
&= \rho \ln\left[\frac{M_{ij,t-1}}{L_{i,t-1}L_{j,t-1}}\right] \\
&+ g\left(\ln\left[W_{jt}/W_{it}\right]\right) \\
&+ \beta_1 \ln\left[u_{jt}/u_{it}\right] + \beta_2 \ln\left[HPI_{jt}/HPI_{it}\right] + \beta_3 \ln\left[age_{jt}/age_{it}\right] \\
&+ \alpha_{ij} + \lambda_t + \varepsilon_{ijt}
\end{aligned}
\tag{7}
$$

where,

$\ln\left[\frac{M_{ijt}}{L_{it}L_{jt}}\right]$: natural log of the population migration flow normalized by the labor forces in destination *j* (*1, ..., N*) and origin *i* (*1, ..., N*) at time *t (1, ..., T)*; $\ln\left[W_{jt}/W_{it}\right]$: natural log of wage differential between *j* and *i* at time t;

$\ln\left[u_{jt}/u_{it}\right]$: natural log of unemployment rates differential between *j* and *i* at time t;

$\ln\left[HPI_{jt}/HPI_{it}\right]$: natural log of housing price index differential between *j* and *i* at time t;

$\ln\left[age_{jt}/age_{it}\right]$: natural log of the proportion of population aged 25-34 (35-44) in region *i* at time t;

ε_{ij} : i.i.d. error term.

We also include spatial fixed effects (α_{ij}) in order to capture the effect of unobservable or omitted variables related to amenities, preferences, social conditions, distance (representing a fixed cost of moving) and so on. Ignoring unobserved location-specific effects is likely to result in biased parameter estimates since these effects must be expected to be correlated with the observed explanatory variables. Finally, we include temporally-specific effects (λ_t) in order to control for the temporal variation of national economic conditions.

$g\left(\ln\left[W_{jt}/W_{it}\right]\right)$ is an unknown function. We only allow the wage variable to make up the non-linear part of the model, while all the other variables enter the model linearly. We use a penalized cubic regression spline to estimate $\hat{g}\left(\ln\left[W_{jt}/W_{it}\right]\right)$. In particular, we apply the method described in Wood (2001) and Wood and Augustin (2002) that allows integrated smoothing parameter selection via GCV (Generalized Cross Validation). This method (implemented in the R package *mgcv*) helps overcome the difficulties of model selection typical of the additive model framework based on back-fitting developed by Hastie and Tibshirani (1990).

Figure 3 shows the fitted smooth function $g\left(\ln\left[W_{jt}/W_{it}\right]\right)$ alongside Bayesian confidence intervals (see Wood, 2004). The vertical axis reports the scale of the expected values of the log of regional migration rate; the horizontal axis reports the scale of the log of interregional wage differentials. A simple F test suggests a significant effect of wages (the F statistic is 11.30 with a p-value of 0.000).[8] Moreover, the result of the specification test for the null hypothesis of a linear model against the semi-parametric alternative suggests that the null hypothesis can be rejected at the 1% level (F = 19.20 with a p-value =0.000).

<<insert Figure 3 here>>

In the interpretation of the result shown in figure 3, it is useful to partition the graph into two parts, one located on the left side of the value of 0.0 on the horizontal axis, and the other on the

[8] The F test in a nonparametric estimation has the same meaning of the F test for the evaluation of the explicative power of each independent variable in the linear regression models.

right side of the value of 0.0 on the same axis. For the former, wage levels in the origin places are higher than those in the destination, while for the latter, the reverse is the case. In the left part, there is no significant relationship between wage differentials and population migration flows, since the confidence interval is quite large and contains the horizontal axis, where no migration is expected. In the right part, where the wage level in the destination exceeds that in the origin, our model still does not predict any movement of population up to a certain threshold (about 0.25). Beyond that threshold, migration is expected to occur and to increase steeply with increasing wage differentials.

This result strongly corroborates the 'option value of waiting' hypothesis of a non-linear relationship between wages and population migration described in section 3. In particular, while the traditional economic theory of migration postulates that workers decide to move when the wage differential exceeds the fixed cost of moving, our analysis on the U.S. case confirms the 'option value of waiting' theory, according to which individuals do not migrate until the wage differential rises substantially above a certain threshold.

Table 1 shows regression results of the semi-parametric formulation. All the coefficients are significant at the 1% level and with the expected sign, except for the variable "proportion of population aged 25-34". The evidence of a negative effect of the log differential of unemployment rates is perfectly coherent with the results of previous analyses on interregional migration. A lower unemployment rate in the destination or a higher unemployment rate in the origin encourages people to migrate.

The effect on the migration rate of the variable measuring the percentage of population in the age class from 35 to 44 in the origin is positive, significant and very large in magnitude, suggesting that the age structure of the population plays a very important role in determining interregional migration patterns in the U.S.

The model also includes a measure of the housing price index in order to capture the effect of the differences in the cost of living on interregional migration. The coefficient of this variable is negative and significant at the 1% probability level. This result is not surprising since the indicator of wage differential included in the model is measured in current prices and it is not corrected for the interregional differential in the cost of living.

As mentioned above, spatial fixed effects are included in the model to reflect the influence of regional amenities and fixed costs of moving. Most of the estimated coefficients α_{ij} are significantly different from zero indicating that there is heterogeneity among regions that cannot be controlled for by the variables included in the model. Finally, the lag term of the dependent variable is significant at the 1% level with a parameter of 0.186 corroborating the hypothesis of persistence in migration behavior.

In order to capture the non-linearity shown in figure 4, we apply a polynomial transformation of the model in $\ln\left[W_{jt}/W_{it}\right]$. We experimented with a fifth-, forth- and third-degree polynomial specification and found that a cubic polynomial fit performs quite well:

$$
\begin{aligned}
y_{ijt} &= \ln\left[\frac{M_{ijt}}{L_{it}L_{jt}}\right] \\
&= \rho y_{ij,t-1} \\
&+\lambda_1 \ln\left[W_{jt}/W_{it}\right] + \lambda_2 \left(\ln\left[W_{jt}/W_{it}\right]\right)^2 + \lambda_3 \ln\left[W_{jt}/W_{it}\right]^3 \\
&+\beta_1 \ln\left[u_{jt}/u_{it}\right] + \beta_2 \ln\left[HPI_{jt}/HPI_{it}\right] + \beta_3 \ln\left[age_{jt}/age_{it}\right] \\
&+\alpha_{ij} + \lambda t + \varepsilon_{ijt}
\end{aligned}
\tag{8}
$$

Indeed, according to an F test, the two models (7) and (8) cannot be considered as statistically different (the F statistic is equal to 0.52 with a p-value of 0.632). The results are reported in the second column of Table 2.

4.2 Econometric results from a GMM estimation of the dynamic migration model

In the second step of the empirical analysis, we estimated the polynomial transformation of the dynamic gravity model (equation 8) using the Generalized Method of Moment (GMM) estimator proposed by Arellano and Bond (1991). As is well known, when the lagged dependent variable is included as a regressor, the within-group (or fixed effects) estimator is biased and inconsistent (even if the other explanatory variables are assumed strictly exogenous) unless the number of time periods is very large (tends towards infinity; on this, see Baltagi, 2005, p. 135). Specifically, the within-group estimate of ρ (the coefficient of the lagged dependent variable) is expected to be downwards-biased because of the negative correlation between the within transformed error term and the within-transformed lagged dependent variable. On the contrary, the GMM estimators are consistent for $N \rightarrow \infty$ and fixed T. The GMM estimators have the further advantage that we do not have to rely on the restrictive assumption of strictly exogenous regressors. In the case of the migration equation, indeed, wages and unemployment levels, as well as living costs and age structure of the population, cannot be considered as strictly exogenous; rather, these variables may be assumed to be predetermined or even endogenous.

The GMM estimator suggested by Arellano and Bond (1991) starts with first differencing the model in (8) in order to eliminate the regional-specific effects α_{ij}. Even if there is no autocorrelation in the model in levels, the error term $\Delta\varepsilon_{ijt}$ of the model in first differences (which will follow an MA(1) process with coefficient -1) is correlated with $\Delta y_{ij,t-1}$ (since $\varepsilon_{ij,t-1}$ is correlated with $y_{ij,t-1}$) implying that we have to use instruments for $\Delta y_{ij,t-1}$ in order to obtain consistent estimates. We may use lagged values of y_{ijt} to form instruments as long as y_{ijt} is lagged two periods or more.[9] If there is first order autocorrelation in the levels equation (8), there will be second order autocorrelation in the first differenced equation implying that $y_{ij,t-2}$ is not a valid instrument (but y_{ijt} lagged three periods or more may be valid). It is therefore very important to test for autocorrelation.

As recommended in Arellano and Bond (1991), lagged levels of y_{ijt} are used as instruments for $\Delta y_{ij,t-1}$: $\left(y_{ij,t^*}, y_{ij,t^*+1}, ..., y_{ij,t-2}\right)$ where t^* is the first year with observations for y_{ijt}.[10] Thus, there are more instruments the higher the value of t. This choice of instruments exploits the moment conditions $\sum_{ij=1}^{N} y_{ijs} \Delta\varepsilon_{ijt} = 0$, $s = t^*, \ldots, t-2$ for each year (t) in the estimation period separately in accordance with the fact that N (the number of individuals) is large whereas T (the number of time periods) is small. The GMM estimation may be viewed as a simultaneous estimation of a system of equations, one for each year, using different instruments in each equation, and restricting the parameters (ρ, β, λ) to be equal across equations (years).

---

[9] The GMM estimator is an extension of the Anderson and Hsiao (1982) estimator (which uses $y_{ij,t-2}$ or $\Delta y_{ij,t-2}$ as instruments for $\Delta y_{ij,t-1}$ in a 2SLS framework) as it exploits additional moment conditions and uses a weighting matrix which takes into account the MA(1) process in $\Delta\varepsilon_{ijt}$ as well as general heteroskedasticity.

[10] The lagged values of y_{ijt} may be weak instruments for $\Delta y_{ij,t-1}$ when the series is highly persistent with a large value of ρ or a large relative variance $\left(\sigma_\alpha^2 / \sigma_\varepsilon^2\right)$ of the individual effects α_{ij} compared to the transitory shocks ε_{ijt} (see Blundell and Bond, 1998, 1999). This problem, which can result in large finite-sample bias of the GMM-difference estimator, may be solved by using the 'System GMM estimator' developed by Blundell and Bond (1998). In our case, however, this problem did not emerge and the results from the GMM-difference estimation appeared as more reliable.

Results from the GMM estimation of equation (8) are shown in table 2 (columns 2, 3 and 4). Long-run parameters with standard errors are reported at the bottom of Table 2.[11] Column 2 of the table shows the (one-step) estimation results for the model specified with strictly exogenous regressors (except for the lagged term of the dependent variable, of course); column 3 reports the (two-step) estimation results for the model specified with predetermined regressors (all instruments are lagged at least one period); in the last model specification (column 4) all variables are treated as endogenous (all instruments are lagged at least two periods). The test statistics of serial correlation (m_1 and m_2) and over-identifying restrictions ('Sargan/Hansen') do not indicate misspecification.[12] Since it may be important to take account the potential endogeneity of the explanatory variables of the model, we consider this last specification as the preferred model.

Broadly speaking, the results of the GMM estimation tend to confirm the 'within group' estimation results (reported in column 1), especially when the covariates are considered as endogenous. As expected, the magnitude of the coefficient of the lagged term of the dependent variable is higher in the GMM estimations. In line with the within-group estimation results, the housing price index differential has a negative effect on population migration. If the housing price index differential is increased by 1%, population immigration will decrease by 0.79% in the short run and by 1.12% in the long run when the variable is treated as endogenous; the negative effect is lower when the variable is treated as exogenous or predetermined.

---

[11] Let β_k denotes the coefficient of a given explanatory variable (the short-run effect of this variable). The long-run effect is equal to $\beta_k/(1-\rho)$, where ρ is the coefficient of the lagged dependent variable. Standard errors of long-run parameters are calculated using the Delta method (see Greene, 1997).

[12] The test statistics m_1 and m_2 test for presence of serial correlation in the first differenced residuals of first and second order, respectively; they are asymptotically normally $N(0,1)$ distributed under the null of no serial correlation (see Arellano and Bond, 1991). The results show that there is no significant second order autocorrelation which is the crucial point with respect to the validity of the instruments. The Sargan test statistic of overidentifying restrictions is χ^2-distributed with degrees of freedom equal to the number of instruments minus the number of estimated parameters. This misspecification test does not indicate correlation between the instruments and the error term.

The “proportion of population aged 25-34” turns out to be insignificant in the fixed effect and in the GMM with predetermined or endogenous variables, while it is significant and positive in the GMM with strictly exogenous variables. The “proportion of population aged 35-44” is instead positive and significant even it is treated as endogenous. The short run effect of this variable is 2.89, while the long run elasticity is 4.10. The unemployment rate enter significantly in the cases of within-group and GMM with endogenous variables: an increase of the unemployment rate differential by 1 percentage point leads to a decrease in population immigration of 0.41% in the short run and of 0.59% in the long run.

However, the most important result, from our point of view, is the evidence of a non-linear relationship between population migration and wage differentials, confirming the ‘option value of waiting’ theory. In fact, the quadratic and cubic terms of wage differential are always statistically significant and positive (except for the quadratic term in column 2). The magnitude of their coefficients in the within-group and in the GMM with endogenous variables does not differ dramatically. In particular, when wage differentials are treated as endogenous variables (column 4), the short run effect of the quadratic term is 1.01 while its long run effect is 1.43; the short run effect of the cubic term is 2.50 while its long run elasticity is 3.55. In other words, the results of the GMM panel estimations strongly confirm the semi-parametric evidence discussed in the previous section and, thus, corroborate the ‘option value of waiting’ hypothesis.

4.3 Detecting spatial patterns

In the econometric analysis carried out in this version of the paper we do not take into account the possible presence of spatial dependence.[13] Here, however, we ask whether there is any clear spatial pattern in the error term of the model. Most of the errors found in the proposed model appear in the origin-destination pairs which have a Midwestern MSA either as origin or destination and Midwestern MSAs as both origins and destinations in some cases. This can be explained by two aspects. First, relatively smaller population size of Midwestern MSAs included, especially for Cleveland, OH, Cincinnati, OH-KY-IN, Indianapolis, IN, Milwaukee-Waukesha, WI, may cause different interregional migration flow patterns compared to the non-Midwestern MSAs which are generally larger in population size. Secondly, geographic structure of MSAs included in this study may cause the abnormal interregional migration flow patterns of certain origin-destination pairs which cannot be effectively explained by the proposed model. For example, interregional migration flows among closely located Midwestern MSAs may have other motivations not related to the economic utility maximization. Consequently, other types of explanations may be more appropriate to describe the shorter moves among Midwestern MSAs.

5. Conclusions

Traditionally, migration decision process is believed to be composed of two criteria: “whether to move” and “where to move”. However, the possible existence of ‘option value of waiting’ in interregional migration adds another decision criterion, “when to move”. Population at-risk can be regarded as potential migrants from an origin place. For the potential migrants, they move to maximize their utilities based on destination choice. With the ‘option value of waiting,’ potential

[13] Only recently, there has been a first attempt to extend the gravity model to explicitly take into account spatial dependence in the origin-destination flows (see LeSage and Pace, 2005).

migrants can enhance their utilities by postponing their decision to move. Because of the 'option value of waiting', a potential migrant actually moves only when the wage differential between origin and destination places exceeds a certain threshold, which might be much higher than the Marshallian trigger.

In this paper, we exploit the panel structure of a dataset on interregional migration among nineteen MSAs in the US from 1993 to 2001 to test the prediction of a non-linear relationship between interregional migration and wage differentials. In the first step of the analysis, we estimate a dynamic gravity migration model using semi-parametric estimators for an empirical test. The results of this analysis clearly suggest that with a wage differential smaller than a certain threshold, people rarely move controlling for the other socioeconomic variables. Only beyond the threshold, the interregional migration grows rapidly proving an important role of the option value of waiting in migration decision process. In the second step of the analysis, we use GMM estimators in order to take into account possible endogeneity of the explanatory variables. The results of this further analysis confirm the semi-parametric evidence of non-linearity between interregional migration and wage differentials and, thus, proves the robustness of the results obtained in the first step.

Essentially, our study suggests that when a local region suffers from net outflow of population, this may not be solely explained by the wage differentials, especially when this differentials are not big enough to overcome the wage-threshold connected to the option value of waiting. Rather, other socioeconomic factors, such as unemployment differentials and housing price differentials, may have direct impact on decision making process for interregional migration. This has some important implications for regional development policy. In particular, our analysis implies that if a local government is willing to invite more labor to its region for fast growing economy, the policy regarding the stabilization of local housing price and of unemployment, would be the more effective approach rather than directly intervene into the labor market with the purpose of controlling wage levels.

Final considerations regard future challenges of this analysis. As explained in the previous section, interregional migration flows among certain pairs of origin-destination cannot be effectively explained by the proposed model. Considering urban hierarchy in terms of population size and spatial structure of certain regions included in our study, future studies should address the issues related to the hierarchy and spatial structure. Controlling the urban hierarchy and spatial structure among the regions will enable us to propose more strong empirical evidence on the role of wage differentials with option value of waiting in the migration decision making process.

References

Anderson, T.W. and Hsiao, C. (1982), Formulation and estimation of dynamic models using panel data. Journal of Econometrics 18, 47–82.

Arellano, M. and Bond, S. (1991), Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. Review of Economic Studies 58, 277– 297.

Baltagi, B.H. (1995), Econometric Analysis of Panel Data. Wiley, Chichester.

Bauer, T. and K. Zimmermann (1995), "Integrating the East: the labour market effects of immigration", *Centre for Economic Policy Research Discussion paper* no.1235.

Burda, M.C. (1993), "The determinants of East-West German migration", *European Economic Review*, 37, 452-461.

Blundell, R. and S. Bond (1998), "Initial Conditions and Moment Restrictions in Dynamic Panel Data Models", *Journal of Econometrics*, 87, 115-143.

Blundell, R. and Bond, S. (1999), GMM estimation with persistent panel data: an application to production functions Institute for Fiscal Studies. Working paper series no. W99/4.

Decressin, J. W. (1994), "Internal Migration in West Germany and Implications for East-West Salary Convergence", *Weltwirtschaftliches Archiv*, 130, 231-257.

Ghatak, S., P. Levine and P.S. Wheatley (1996), "Migration theories and evidence: an assessment", *Journal of Economic Surveys*, 10, 159-198.

Gordon, I. (1985), "The Cyclical Interaction Between Regional Migration, Employment and Unemployment: A Time-Series Analysis for Scotland", *Scottish Journal of Political Economy*, 32: 135-158.

Greene, W.H., 1997. Econometric Analysis. Prentice Hall, New Jersey.

Greenwood, M.J. (1997), "Internal Migration in Developed Countries," in M.R. Rosenzweig and O. Stark (eds.), *Handbook of Population and Family Economics*. Amsterdam: Elsevier, 647-720.

Greenwood, M.J. and G. Hunt (2003), "The Early History of Migration Research", *International Regional Science Review*, 26(1): 3-37.

Harris, J., and M. Todaro (1970), "Migration, Unemployment, and Development: A Two-Sector Analysis", *American Economic Review*, 60: 126-142.

Hart, R. (1975), "Inter-Regional Migration: Some Theoretical Consideration (Part 2)", *Journal of Regional Science*, 15(3): 289-305.

Hastie and Tibshirani (1990) Generalized Additive Models. Chapman and Hall.

Hicks, J. (1932), *The Theory of Wages*, London: Macmillan.

LeSage, J. and R. K. Pace (2005), Spatial Econometric Modeling of Origin-Destination Flows, mimeo, University of Toledo.

Makower, H.J., J. Marschak and H.W. Robinson (1938), "Studies in Mobility of Labor: A Tentative Statistical Measure", *Oxford Economic Papers*, 1: 83-123.

Makower, H.J., J. Marschak and H.W. Robinson (1939), "Studies in Mobility of Labor: Analysis for Great Britain, Part I", *Oxford Economic Papers*, 2: 70-97.

Makower, H.J., J. Marschak and H.W. Robinson (1940), "Studies in Mobility of Labor: Analysis for Great Britain, Part II", *Oxford Economic Papers*, 4: 39-62.

Molho, I. (1984), "A Dynamic Model of Inter-Regional Migration Flows in Great Britain", *Journal of Regional Science*, 24: 317-337.

Molho, I. (1986), "Theories of Migration: A Review", *Scottish Journal of Political Economy*, 33: 396-419.

Oswald, A.J. (1990), *The Housing Market and Europe's Unemployment: A Non-technical Paper*, mimeo University of Warwick.

Parikh, A. and M. Van Leuvensteijn (2002), "Internal migration in regions of Germany: a panel data analysis", *ENEPRI Working Paper* no. 12.

Pissarides, C.A. and I. McMaster (1990), "Regional Migration, Wages and Unemployment: Empirical Evidence and Implications for Policy", *Oxford Economic Papers,* 42, 812-831.

Plane, D.A., and P.A. Rogerson (1994), *The Geographical Analysis of Population: With Applications to Planning and Business*, (John Wiley & Sons, New York, NY).

Rogers, A. (1990), "Requiem for the Net Migrant", *Geographical Analysis*, 22: 283-300.

Rogers, A., and L. Castro (1986), "Migration", in A. Rogers and F.J. Willekens (eds.) *Migration and Settlement: a Multiregional Comparative Study*, (D. Reidel, Boston), pp. 157-208.

Rogers, A., R. Racquillet, and L. Castro, (1978), "Model Migration Schedules and Their Applications", *Environment and Planning A*, 10: 475-502.

Sjaastad, L. (1962), "The Costs and Returns of Human Migration", *Journal of Political Economy*, 70: 80-93.
Todaro, M.P. (1969), "A Model of Labor Migration and Urban Unemployment in Less Developed Countries", *American Economic Review*, 59: 138-148.
Todaro, M.P. (1976), *Internal Migration in Developing Countries: A Review of Theory, Evidence, Methodology, and Research Priorities*, International Labor Office, Geneva.
Wood (2001) mgcv: GAMs and Generalized Ridge Regression for R. R News 1(2):20-25.
Wood, S.N. (2004) On confidence intervals for GAMs based on penalized regression splines. Technical Report 04-12 Department of Statistics, University of Glasgow.
Wood and Augustin (2002) GAMs with integrated model selection using penalized regression splines and applications to environmental modelling. Ecological Modelling 157:157-177.

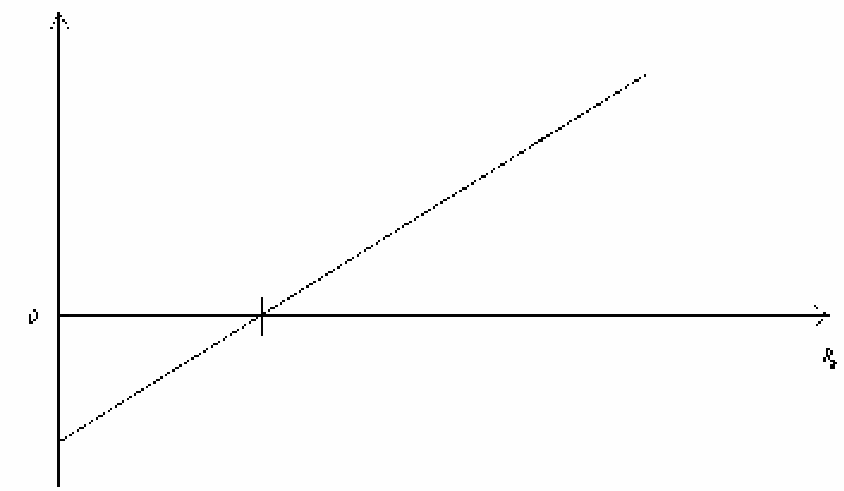

Fig. 1 The migration decision under certainty

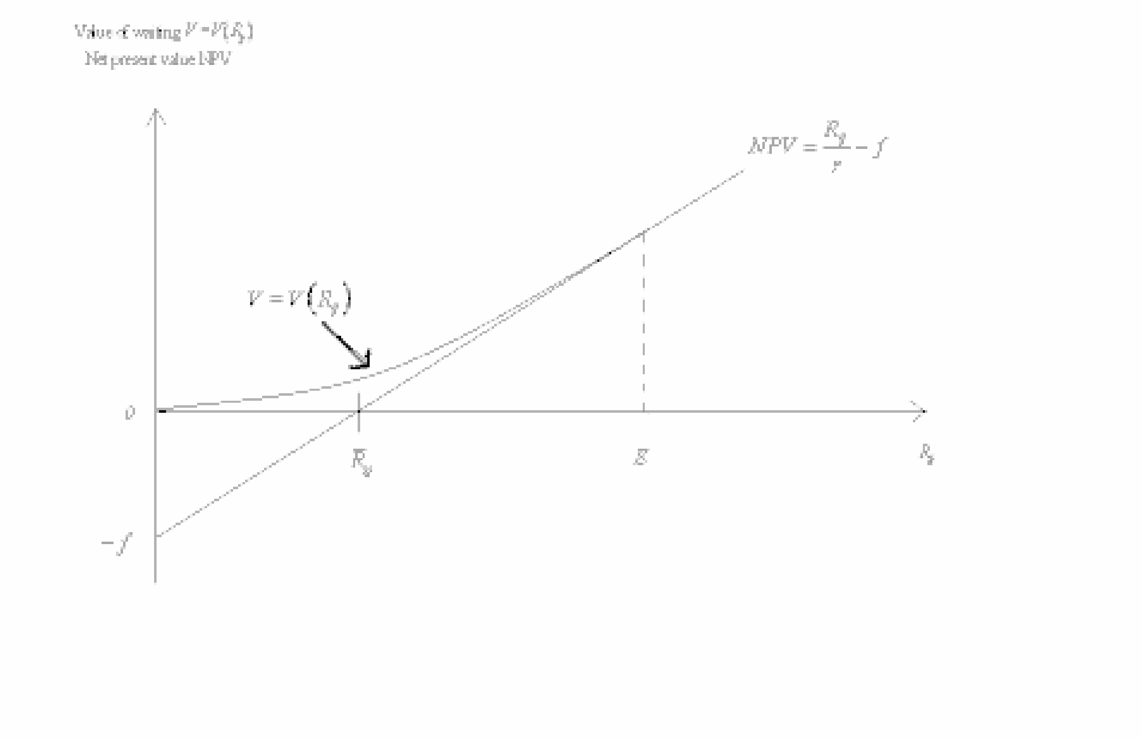

Fig. 2 The migration decision under uncertainty and the option of waiting

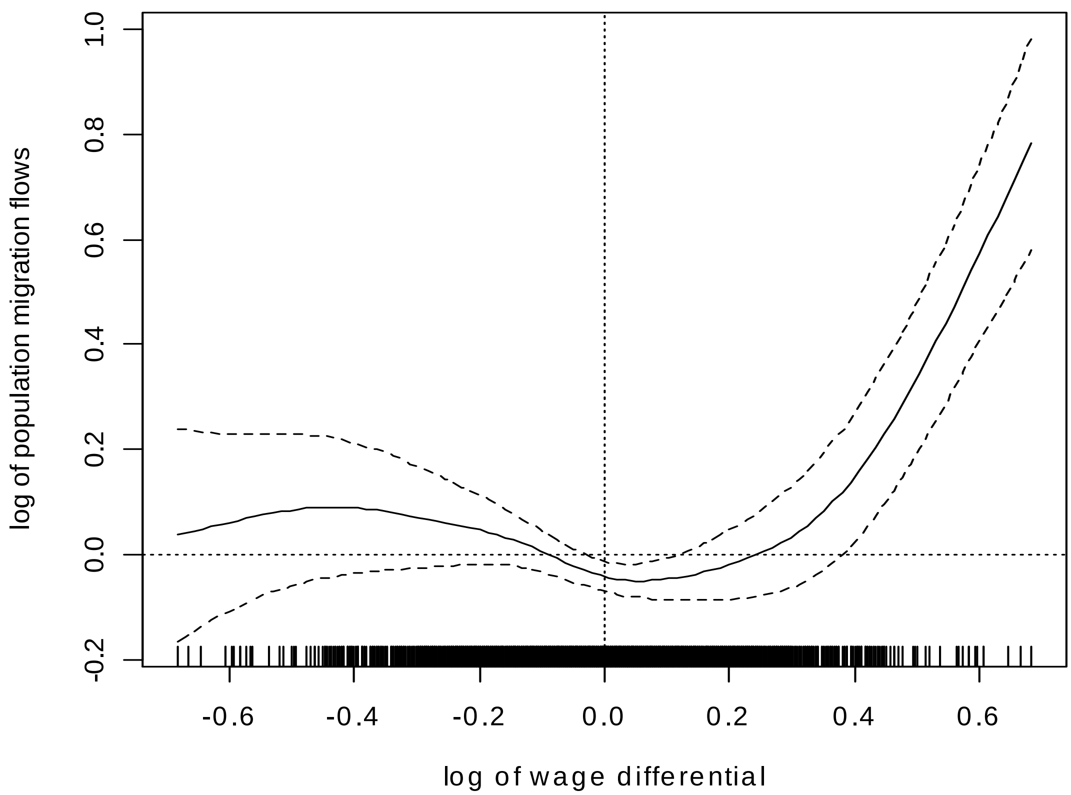

Figure 3 – Plot of the estimated partial regression function for the additive regression of migration on wages. The partial regression uses a spline smother. The broken line gives pointwise 95% confidence envelopes for the partial fit. Note: Approximate significance of smooth term $g\left(\ln\left[W_{jt}/W_{it}\right]\right)$*: F test = 11.30 (p-value=0.000).*

Table 1 – Estimation results for the dynamic gravity migration model 1994-2001.

Semi-parametric regression

	Semi-parametric (fixed effects)
Lag of the dependent variable	0.288 (0.000)
Wage differential	See figure 3
Unemployment rate differential	-0.118 (0.001)
House price index differential	-0.557 (0.000)
Proportion of population aged 25-34	0.159 (0.294)
Proportion of population aged 35-44	1.699 (0.000)
N	2720

Notes: All regressions are estimated in R 2.2.0. All estimates include a full set of time dummies as regressors. The constant term is excluded. P-values are in round brackets.

Table 2 – Estimation results for the dynamic gravity migration model 1994-2001.

Fixed effects 3-degree polynomial regression and First difference GMM

	Within-Group	**GMM-Diff** Strictly exogenous regressors (one-step)	**GMM-Diff** Predetermined Regressors (two-step)	**GMM-Diff** Endogenous wages (two-step)
	1	**2**	**3**	**4**
Lag of the dependent variable	0.287 (0.000)	0.380 (0.000)	0.380 (0.000)	0.295 (0.000)
Wage differential	-0.339 (0.032)	-0.610 (0.017)	-0.397 (0.431)	-0.603 (0.212)
Square of wage differential	0.999 (0.000)	0.176 (0.552)	0.638 (0.009)	1.010 (0.000)
Cube of wage differential	2.062 (0.000)	1.427 (0.016)	2.465 (0.001)	2.502 (0.000)
Unemployment rate differential	-0.119 (0.001)	0.014 (0.797)	-0.151 (0.172)	-0.413 (0.000)
House price index differential	-0.537 (0.000)	-0.238 (0.006)	-0.614 (0.001)	-0.791 (0.000)
Proportion of population aged 25-34	0.451 (0.140)	2.000 (0.000)	-3.356 (0.169)	1.437 (0.416)
Proportion of population aged 35-44	1.423 (0.003)	4.814 (0.000)	2.373 (0.248)	2.891 (0.068)
N	2720	2380	2380	2380
Sargan/Hansen test		28.12 (0.211)	40.56 (0.142)	50.06 (0.110)
m_1		-8.14 (0.000)	-5.95 (0.000)	-7.56 (0.000)
m_2		1.20 (0.231)	-0.47 (0.636)	-0.35 (0.727)
LONG-RUN COEFFICIENTS				
Wage differential	-0.476 (0.031)	-0.984 (0.028)	-0.641 (0.413)	-0.856 (0.192)
Square of wage differential	1.403 (0.000)	0.284 (0.556)	1.031 (0.006)	1.433 (0.000)
Cube of wage differential	2.893 (0.000)	2.303 (0.015)	3.979 (0.000)	3.550 (0.000)
Unemployment rate differential	-0.167 (0.001)	0.023 (0.798)	-0.245 (0.184)	-0.587 (0.000)
House price index differential	-0.754 (0.000)	-0.384 (0.004)	-0.992 (0.002)	-1.121 (0.000)
Proportion of population aged 25-34	0.633 (0.140)	3.227 (0.000)	-5.417 (0.146)	2.039 (0.428)
Proportion of population aged 35-44	2.000 (0.004)	7.767 (0.000)	3.830 (0.245)	4.102 (0.062)

Notes: All regressions are estimated in Stata 9.0. WG estimates include a full set of time dummies as regressors. GMM estimates include a full set of time dummies, regional dummies (indicating, respectively, Non-Midwest to Non-Midwest, Non-Midwest to Midwest, Midwest to Non-Midwest, Midwest to Midwest) and interactions between time dummies and regional dummies as regressors and instruments for the equations in differences. The constant term is always excluded. The null hypothesis that each coefficient is equal to zero is tested using one-step robust standard errors in the case of exogenous variables and two-step standard errors in the other two cases. $m_1(m_2)$ is a test of the null hypothesis of no first (second) order serial correlation, while the Sargan/Hansen test is a test for the validity of the overidentifing restrictions. P-values are in round brackets. Sargan and $m_1(m_2)$ tests are always from the two-step estimation.

Appendix

- List of MSAs included in the model and their locations

Code	MSA Name		
520	Atlanta, GA	3480*	Indianapolis, IN
1120	Boston, MA-NH	4480	Los-Angeles-Long Beach, CA
1600*	Chicago, IL	5080*	Milwaukee-Waukesha, WI
1640*	Cincinnati, OH-KY-IN	5120*	Minneapolis-St. Paul, MN-WI
1680*	Cleveland-Lorain-Elyria, OH	5600	New York, NY
1840*	Columbus, OH	6160	Philadelphia, PA-NJ
1920	Dallas, TX	7040*	St. Louis, IL-MO
2160*	Detroit, MI	7360	San Francisco, CA
3360	Houston, TX	7400	San Jose, CA
		8840	Washington, DC-MD-VA-WV