Measuring and Interpreting Spatial Interactions in Population Migration: An Example of Interprovincial Migration Flows Based on the 6th and 7th Population Censuses

LU Xigang, SONG Zuxuan, PANG Lei

Abstract: Migration flows are determined by both place-specific attributes and spatial interactions between origin and destination locations. However, conventional gravity models often struggle to disentangle these effects—specifically, the spatial interaction component, which is typically proxied by distance and remains the most challenging aspect to model, thereby hindering theoretical interpretation. Drawing on migration flow decomposition and structural gravity model theory, this study replaces the one-step "flow-variable" approach with a two-step "flow-effect-variable" framework. Using interprovincial migration data from China's 6th and 7th population censuses, we decompose migration flows into spatial influence (origin), attractiveness (destination), and spatial interaction components to analyze their temporal trends. Building on these metrics, we employ multiple regression analysis to interpret the factors shaping spatial interactions, with a specific focus on the relationship between bilateral explanatory variables and spatial structure effects.

Keywords: spatial interaction; migration flow decomposition; spatial structure effect; interprovincial migration; 6th population census; 7th population census

1. Introduction

Spatial interaction is a core theme in spatial disciplines. Gravity models are widely applied to analyze various flows—such as migration, commuting, trade, capital, and knowledge—across geographical spaces. However, the relationship between "flows" and "spatial interactions" remains ambiguous, as flows are often used as a proxy for the interactions themselves. Generally, migration flows are the product of three distinct forces: the push capacity of the origin (out-migration potential), the pull capacity of the destination (in-migration attractiveness), and the spatial interaction between the two.

Conventional gravity models treat migration as a "black box," modeling flows as a direct function of explanatory variables. While adding more variables may improve model fit, this approach faces three major challenges:**Measurement of Effects:** Explanatory variables are merely approximations rather than the effects themselves.

Theoretical Interpretation: By conflating the two distinct processes of "variable-to-effect" and "effect-to-flow" into a single step, these models struggle to identify the mechanisms behind migration or the sources of model bias.

Spatial Structure Effects: While spatial econometric models exist to address spatial autocorrelation, they often fail to integrate spatial structure effects into the endogenous logic of spatial interaction models.To address these, this study adopts a flow decomposition approach to measure push, pull, and spatial interaction effects independently. Focusing on the most significant source of error in empirical models—spatial interaction—we utilize data from China's 6th and 7th population censuses to measure and interpret these patterns.

1.1 Migration Theory and Spatial Interaction Models

Research on migration dynamics generally follows two complementary paradigms: theoretical explanation and model construction.

Theoretical explanations emphasize either equilibrium or disequilibrium factors. Early theories highlighted regional economic disparities (e.g., wage gaps, employment opportunities) as drivers. However, empirical discrepancies led to the "spatial utility" interpretation, suggesting that non-economic factors like public services, social atmosphere, and environmental comfort compensate for economic differences. Furthermore, structural theories, such as World Systems theory and Social Network theory, emphasize the self-sustaining nature of migration flows.

Migration models focus on spatial patterns but have struggled with the "representation of spatial interaction." While models have evolved from simple gravity equations to complex formulations, they still predominantly rely on distance-based functions. Because distance is an incomplete proxy—failing to capture psychological, cultural, or institutional costs—this leads to "distance puzzles," "border puzzles," and non-monotonic decay. Consequently, spatial structure effects arise from the omission of these variables, yet they remain exogenous to the core interaction model.

1.2 Migration Flow Decomposition

The decomposition method, tracing back to Shryock and further developed by Mueser and Rogers et al., allows for the analytical separation of migration flows into generation, attraction, and interaction components. Unlike external approximation methods, statistical decomposition (using multiplicative or log-linear models) provides an endogenous measure derived directly from the observed flow matrix. Furthermore, structural gravity models in trade research provide a robust theoretical basis for this decomposition.

2. Data and Methods

2.1 Research Data The study utilizes data on 930 interprovincial migration flows among 31 provincial-level administrative units (excluding Hong Kong, Macau, and Taiwan) from the 6th and 7th censuses. The definition of interprovincial migrants follows the international statistical standard: residents whose current place of residence differs from that of five years prior.

2.2 Research Methods

2.2.1 Migration Flow Decomposition

We utilize a multiplicative component model where a migration flow (M_{ij}) is represented as:

$M_{ij} = T \cdot O_i \cdot D_j \cdot F_{ij}$

Where:T is the total effect (system-wide migration volume).

O_i is the origin (push) capacity.

D_j is the destination (pull/attractiveness) capacity.

F_{ij} is the spatial interaction effect.This framework allows us to define the spatial interaction effect as the deviation of observed flows from the expected values under the assumption of "no spatial interaction" (a state dependent solely on total volume and origin/destination capacities).

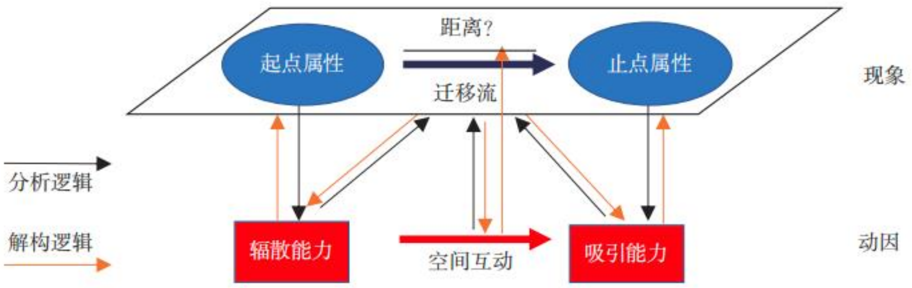

Fig.1 Schematic diagram of migration flow structure

2.2.2 Interpretation of Spatial Interaction

To interpret F_{ij}, we incorporate five bilateral variables:**Distance (DS):** Railway travel time between provincial capitals.

Border (BL): Shared border length.

Language Distance (LD): Weighted index based on dialect and language family mapping.

Migration Network (MN): Ratio of migrant stock from the previous census to the resident population, reflecting temporal lag and network effects.

Spatial Structure Effect (SE): Following Baier and Bergstrand, we incorporate "multilateral resistance" (MR) to account for the relative nature of spatial interactions.3. Findings and Discussion

3.1 Spatial Interaction Patterns The data reveals that while total interprovincial migration volume peaked during the 6th census and slightly declined by the 7th, the spatial pattern shifted from polarization toward equilibrium. Our decomposition shows that the "spatial interaction" effect exhibits a fundamentally different pattern than the raw migration flows. While raw flows are dominated by the push/pull effects of the central-western to coastal shift, high-intensity spatial interactions are clustered in peripheral regions (Northwest, Northeast, South, and Southwest).

3.2 Impact of Spatial Interactions Regression analysis confirms that destination attractiveness remains the primary driver of Chinese internal migration, followed by origin push capacity. Spatial interaction has the lowest explanatory power among the three, yet it remains the most significant source of model error in conventional studies due to its poor representation. Furthermore, the stability of spatial interaction effects over time is less robust than that of pull effects, challenging the assumption of temporal stability in spatial structure.

3.3 Modeling and Interpretation Our models indicate that incorporating multilateral resistance (MR) significantly improves model fit (from $R^2 = 0.397$ to 0.778). The shift from absolute bilateral variables to "relative" metrics (incorporating MR) clarifies that spatial interaction is not merely an absolute cost between two points, but a relative cost within the entire national system. Notably, the significance of migration networks remains high, suggesting that migration has transitioned from being driven by initial friction (distance/language) to being sustained by self-reinforcing social networks.

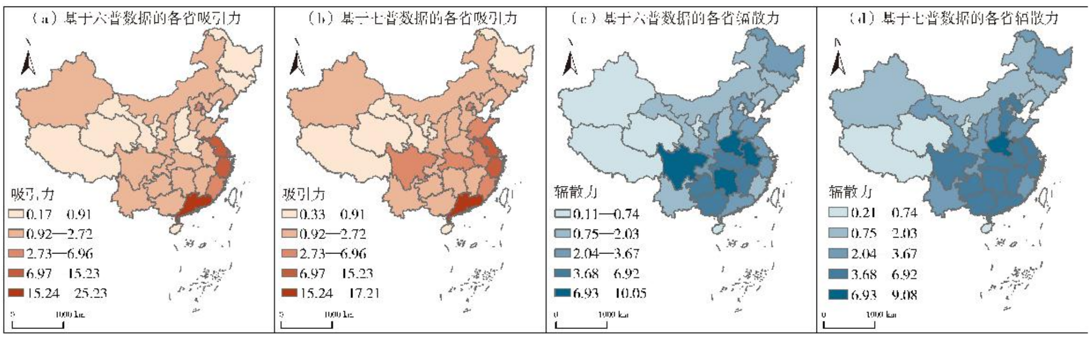

Fig.2 Attraction and spatial influence of provinces based on the 6th and the 7th Population Censuses

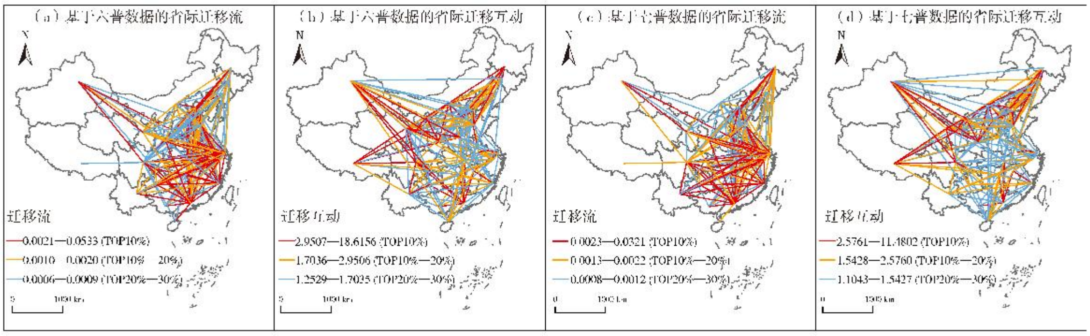

Fig. 3 The top 30% migration flow and its spatial interaction based on the 6th and the 7th Population Censuses

4. Conclusion

This study bridges the gap between migration theory and empirical modeling. By introducing a two-step "flow-effect-variable" framework, we demonstrate that spatial structure effects are essentially the mediating relationship between bilateral absolute variables and their relative system-wide counterparts. Our findings highlight that migration patterns in China are increasingly shaped by self-sustaining network effects, suggesting that policy interventions should move beyond traditional distance-based considerations to focus on the structural maintenance of migrant networks.

（AI 辅助翻译）