The impact of cities' transportation network connections on regional market integration: the case of China's urban agglomerations

Abstract

Despite growing scholarly attention on the role of urban networks for understanding regional dynamics, there has been limited research examining the impact of cities' transportation network connections on regional market integration. This paper analyzes China's four major urban agglomerations: the Yangtze River Delta, the Pearl River Delta, Beijing-Tianjin-Hebei, and Chengdu-Chongqing. Applying a spatial Durbin model to cross-sectional datasets for 2019, we provide insight into the role of cities' transportation network connections in promoting regional market integration, considering both the potentially heterogeneous impact of network connections and the interplay between network and agglomeration externalities. Our results indicate that: (1) cities' transportation network connections have an inverted 'U'-shaped effect on regional market integration; (2) transportation network connections have spatial spillover effects; (3) the positive impact of transportation network connections on regional market integration becomes more pronounced as city size decreases; and (4) there are neither complementary nor substitution effects between network and agglomeration externalities. We reflect on the broader implications of our empirical findings for regional development strategies and discuss possible avenues for further research.

Keywords: transportation network connections; network externalities; agglomeration externalities; regional market integration; urban agglomerations; China

Statements and Declarations

Competing Interests:

The authors declare that they have no competing interests, financial or non-financial, that are directly or indirectly related to the work submitted for publication.

Data Availability

The datasets generated during and/or analyzed during the current study are available from the corresponding author upon reasonable request.

1 Introduction

Cities are increasingly understood as both agglomerations and nodes in networks (Castells, 2000; Bathelt et al., 2004; Amin & Thrift, 1992). The latter has crystallized in an emerging 'urban network paradigm' that has become a significant field of study in urban and regional research. City networks are inherently multiplex: they comprise multifunctional, multidirectional, and multiscalar flows between cities (Burger & Meijers, 2016; Burger et al., 2014). In this context, it is sometimes argued that regions' performance (in the broadest sense of that word) increasingly depends on their network connections (Huang et al., 2020). Earlier research has looked at the importance of cities' network connections for understanding their innovation capabilities (Cao et al., 2022), metropolitan functions (Meijers et al., 2016), and economic growth (Huang et al., 2020). In addition, there has been research looking at the impact of city connectivity on production, distribution, and consumption patterns (Johansson & Quigley, 2004), as it has the potential to (1) increase the diversity of products and the access to a broader market (Meijers et al., 2016; Wei et al., 2022), (2) intensify long-distance market connections, cooperation and competition on a larger scale (Capello, 2000; Shi et al., 2019), and (3) leverage diverse resources flows of labor, capital, commodities, and information (Cheng & LeGates, 2018; Tian et al., 2022).

Despite these attempts to link urban networks and market activities, there has – to the best of our knowledge – been no research focusing on the impact of cities' transportation network connections on regional market integration (RMI). RMI refers to an evolution towards both less restricted access to markets and the free movement of goods, services, and capital (Poncet, 2004). The manifestation of RMI becomes evident as geographically separated markets interconnect, where price changes are transmitted between markets and tend to uniformity (Fossati et al., 2007). In other words, price relationships are associated with RMI (Ke, 2015; Su et al., 2021), defined here as price convergence in regional markets (Dawson & Dey, 2002; Jacks, 2005). Analyzing how transportation networks and RMI are connected is of significant interest in academic and policy circles. Extant studies argued that transportation networks might promote RMI, as they can enhance market access (Fossati et al., 2007) and facilitate multi-layered flows of population, goods, capital, and information (Lu & Mao, 2020) that allegedly come with lowered transaction costs and smaller price differences (Parsley & Wei, 2001; Niu et al., 2020). However, a systematic analysis is required for a more robust conceptual framing and detailed empirical corroboration of this purported relationship. Furthermore, policy discourses have repeatedly emphasized transportation networks as driving forces for the integration process (Lu & Mao, 2020; Wang et al., 2022). This is particularly relevant in the Chinese context, where policy documents[1] underscore the need to promote RMI through transportation connections. Niu et al. (2020) also echo this policymaking focus, highlighting the importance of linking transportation networks to RMI in territorial-based strategies.

Against this background, we contribute to the literature by examining the relationship between cities' transportation network connections and RMI in China's leading urban agglomerations (UAs). We opt for railway connections to capture transportation networks as this allows us to proximate different types of complementary flows of goods, labor, capital, and knowledge in a single indicator (Jiao et al., 2017; Huang et al., 2020). Data for 2019 is used to mitigate potential biases caused by the Covid-19 pandemic on intercity transportation. Given that market integration is typically reflected in the distribution of tradable goods across markets (Samuelson, 1952), we employ price differences in commodity markets as a proxy for the level of RMI. In doing so, our study extends the state of the art in four main ways: 1) we develop a conceptual framework that systematically relates cities' transportation network connections to RMI; 2) we address potential endogeneity issues through an instrumental variables approach, extending the work of Huang et al. (2020); 3) we use the spatial Durbin model to capture the

[1] Recent examples are the state-orchestrated 'The Yangtze River Delta Urban Cluster Development Plan (2016)' and 'Construction of a Unified National Market (2022)'.

spatial spillover effect, providing evidence of the presence of network externalities; and 4) we discuss the heterogeneous effects of different sizes of cities and measure the influence of transportation networks on RMI and its (potential) interactions with agglomeration externalities.

The remainder of this paper is organized into four sections. Section two presents our conceptual framework, explaining the potential pathways of the impact of transportation networks on RMI, and specifies the four hypotheses guiding the remainder of our research. We then elaborate on the data, variables, and model specifications in Section three. In Section four, we report the main results. Section five concludes our paper with an interpretation and discussion of our most important findings, policy implications, and critical avenues for future research.

2 Literature Review and Hypothesis Development

This section elaborates on our conceptual framework (2.1) and reviews the literature on RMI in China (2.2).

2.1 Conceptual framework and research hypothesis

Our framework consists of three main dimensions (Fig. 1), each of which will be discussed in turn:

1. the presence of spatial spillover effects;
2. the heterogeneous effects of different city sizes; and
3. the interaction between network and agglomeration externalities.

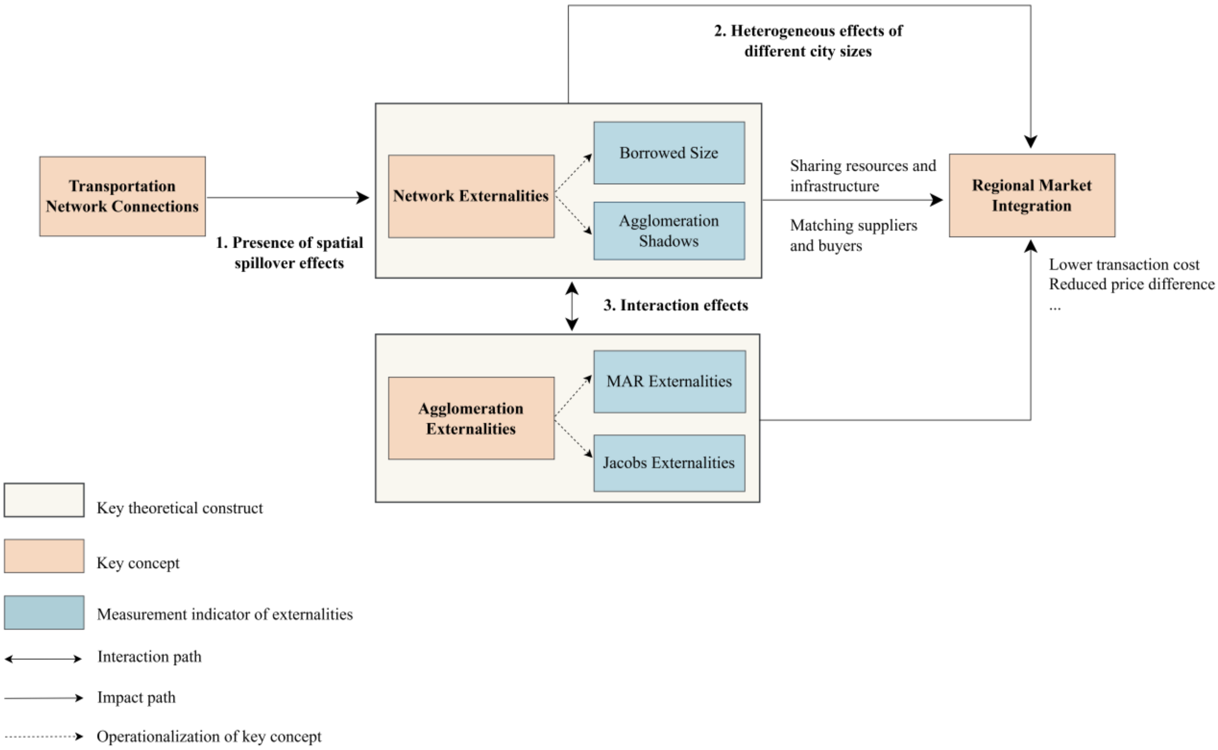

Fig. 1 The conceptual framework of transportation network connections and regional market integration

2.1.1 The interplay between transportation network connections and RMI

When considering the impact of transportation network connections on the level of RMI, the concept of agglomeration externalities is particularly relevant. Agglomeration externalities capture the benefits of being concentrated at a given location (Parr, 2002, p. 152) through sharing, matching, and learning (Duranton & Puga, 2004; Meijers et al., 2016). Against the backdrop of the widening geographical scale of agglomeration benefits, more classical understandings of agglomeration externalities have been extended, among other things, by coining the term 'network externalities' (Capello, 2000). Network externalities refer to the spillovers emanating from the extent to which cities are connected to other cities. As a corollary, the mechanisms of agglomeration externalities are not exclusively confined to cities' 'internal' dynamics but can (also) be shared in networks of cities (Meijers et al., 2016).

The mechanism underlying the hypothesized impact of transportation network connections on RMI can mainly be understood through sharing and matching. First, transportation network connections can share diversified production inputs and suppliers at a larger spatial scale. Cities can also collaborate in sharing infrastructure and technological services (Batten, 1995), such as supply platforms and producer services, which can contribute to shortened transit times, decreased trade costs, and reduced price differentials between cities (Tian et al., 2022). Second, transportation network connections can enhance the movement of people and goods through travel and commuting (Wei et al., 2022) and facilitate information exchange through communication (van Meeteren et al., 2016). Cities can thus benefit from increased opportunities and enhanced flexibility to match supply and demand in regional markets (e.g., matching different types of goods with consumer demands, paring skilled individuals with suitable jobs, etc.) (Niu et al., 2020), and thus contribute to reduced costs and increased levels of integration in regional markets.

However, not unlike the concept of agglomeration diseconomies, transportation network connections can also have a downside: over a certain threshold, they may risk placing a city in a situation where labor, goods, capital, and information are congested (Lu & Mao, 2020). This is empirically supported by Herranz-Loncán (2007) and Cao et al., (2022), who found that when network connections reach a certain level, the benefits of being connected decrease. As the influx of factors brought in by networks exceeds a city's endogenous capacity, the marginal benefits of shared infrastructure and matching supply and demand decrease. This can deliver some adverse outcomes, such as high cost and trade frictions resulting from traffic congestion and information asymmetry, which could result in increased transaction cost and price differences, and impede the market integration process.

Our first hypothesis, therefore, is: *The transportation network connections of a city have an inverted 'U'-shaped effect on its level of RMI.*

2.1.2 The spatial spillover of transportation network connections - network externalities

Burger et al. (2015) expanded our understanding of urban network externalities by recasting it in line with Alonso's (1973) concept of 'borrowed size,' where small cities can borrow some of the functions from neighboring large cities. When the benefits cities gain from their larger neighbors are limited by competition effects, they may experience adverse outcomes: 'agglomeration shadows' (Partridge et al., 2009). The notions of borrowed size, agglomeration shadows, and network externalities can be integrated into a single framework. The former two concepts are akin to different sides of network externalities as they describe the positive and negative spillover effects of being connected in urban networks. Thus, the different outcomes observed within network externalities depend on whether borrowed size or agglomeration shadows dominate. Drawing on data on the distribution of city functions in Western European cities, Meijers et al. (2016) found that agglomeration shadows dominate borrowed size effects, with more cities facing the former than the latter through regional network connectivity. Volgmann and Rusche (2020) and Tian et al. (2022) have recently adopted this theoretical framework, focusing on German urban regions and Chinese cities, respectively.

We argue that these arguments can be extended to understand the impact of transportation connections on RMI, as improvements in regional transportation connectivity can support cities borrowing benefits (Huang et al., 2020), such as high-skilled labor, consumer markets, and trade information through labor movement and information exchanges. This can help reduce transaction costs and regional price differences and thus increase RMI levels. Given this, a city's transportation network connections can generate positive spatial spillovers on the RMI of other cities. However, as connections reach a threshold, the competition process between cities intensifies (Huang et al., 2020), especially for high-skilled labor, production inputs, and consumer demand. This can lead to increased costs, inhibited labor and goods movements (Mačiulis et al., 2009), and increased price differences and market fragmentation in other cities. That is, negative spillover effects – i.e., 'agglomeration shadows' – will become dominant in the regional market.

Our second hypothesis, therefore, is: The transportation network connections of a city have indirect effects on the RMI levels of cities with which it is connected through spatial spillovers.

2.1.3 The heterogeneous influence of transportation network connections

Despite the now often-made claim that externalities can be obtained through network connections, it has been argued that their impact may vary depending on the size of cities. In Alonso's original conceptualization, small cities are more likely to borrow from their larger neighbors. This argument has been supported by Camagni et al. (2016), who found that small cities experience borrowed size effects and gain more productivity through long-distance collaborations than large cities. Therefore, our previous hypothesis that transportation network connectivity can positively impact cities' RMI levels needs to be extended with the hypothesis that these effects may be amplified in smaller cities.

Smaller cities can better internalize the benefits (Burger & Meijers, 2016, p. 6) of being connected to larger cities. More specifically, by being connected to sizeable consumer markets and shared infrastructure, smaller cities can attract different types of production factors and consumption goods (Johansson & Quigley, 2004), contributing to lower transaction costs and higher levels of RMI. Larger cities, in contrast, may experience adverse outcomes because of high transaction costs resulting from increased transit times, higher land prices, and environmental pollution (Tian et al., 2022). These adverse outcomes could hinder the movement of labor and goods, widen the price gap between cities, and amplify the negative effect of transportation network connections on RMI in larger cities. Meanwhile, smaller cities have greater endogenous capacity – lower land cost, less dense population, and improved transportation infrastructure – to mitigate these negative effects (Capello & Camagni, 2000).

Our third hypothesis, therefore, is*: As city size grows, the positive impact of cities' transportation network connections on RMI will decrease, while the negative impact will increase.*

2.1.4 Interactions between agglomeration externalities and network externalities

It is argued that network externalities may complement or substitute for such benefits of agglomeration. Following van Meeteren et al.'s (2016) conceptualization, we attempt to understand the potential interactions between agglomeration and network externalities in trade activities from an industrial relation and a spatial-economic perspective.

First, industrial organization theory claims that stable network relations among firms are the backbone of the economy. An upscaled of this perspective on firm networks has led to an understanding of city networks as a club good (Capello, 2000). Inter-city connections are believed to enhance utilities of participating cities that complement the endogenously created agglomeration externalities (van Meeteren et al., 2016). This potential interaction between both forms of externalities can also apply to RMI regarding the increased utility and reduced transaction cost. For example, Meijers et al. (2016) found that European cities can compensate for their relative lack of metropolitan functions by being connected in transportation networks. Moreover, production and consumption in market transactions can become more diverse through the channel of borrowed size (Johansson & Quigley, 2004), which complements the benefits of agglomeration.

Second, the spatial-economic perspective, in which distance plays an instrumental role, has also been highlighted (van Meeteren et al., 2016). Previous research suggested that network externalities can help extend agglomeration benefits over a larger geographical scale (Meijers et al., 2016). The expanding scale of network externalities can reduce proximity restrictions in regional market transactions. This includes a reduction in the distance-sensitive transaction cost (e.g., transportation cost and information cost) between seller (supplier) and buyer (customer) and in the market price difference (Johansson, 2005). As such, the formation of market supplier-customer links through transportation networks can be a substitute for geographical proximity.

Given this, our fourth and last hypothesis is that: *There is an interaction effect between network externalities and agglomeration externalities affecting cities' RMI.*

2.2 The Chinese context: pleas for increasing market integration in urban agglomerations

Processes of RMI in China have attracted scholarly attention ever since the country's increased integration into world trade (Poncet, 2005). Young (2000) and Poncet (2005) argued that China has gradually evolved into a fragmented market under the control of local governments since the 1980s. The literature has identified several factors that contribute to this fragmentation, the most influential being Young's (2000) observation that the control of local governments would lead to segmented regional markets. Local governments establish trade barriers to protect local industries from being crowded out of the regional market space and maintain sufficient tax revenue (Poncet, 2005; Que et al., 2018). Other factors that have been identified include cities' specific characteristics, such as economic growth, local officials' performance (Fan et al., 2007), and openness to international markets (Niu et al., 2020). In addition, resource abundance has been argued to co-determine local market supply-demand and relative prices difference of goods (Shao & Yang, 2010). The formation of large-scale polycentric urban regions (Liu et al., 2016) has also been found to help reduce transportation costs and regional price divergences (Wang & Wei, 2022).

One of China's most apt specifications of polycentric urban regions is its 'Urban Agglomerations' (UAs), a new spatial framework in China's national policy agenda[2]. But UAs are more than a policy imagination: they are a relevant 'regional' scale of analysis, as these collectively account for over 80% of China's economic output and are established as a normative objective for achieving regional integration in the country's 'new urbanization' policy[3] (Wu, 2016). Morphologically, UAs are defined as regions with densely interconnected cities with complementary economic profiles (Derudder et al., 2022). These regions are new state spaces that increasingly emerge alongside more established administrative divisions, such as provinces in China's regional policymaking (Wu, 2016). UAs are characterized by high population density and intense functional connections and serve as crucial concentrations of industries, labor, and resources for regional market activities (Fang & Yu, 2017). Despite the putative relevance of UAs as a scale of analysis in RMI research, RMI research using UAs as a framework of analysis remains scarce. Existing studies have mainly focused on provinces (see the cases in Ke, 2015; Su & Liang, 2021), highlighting the need for a more nuanced understanding of regional dynamics and their role in RMI.

3 Study area, Data, and Methodology

3.1 Study area

Our study centers on intercity transportation networks as a proxy of the complementary flows of people and information. Our study area deliberately centers on the four poles - the major urban agglomerations (UAs) - identified in China's diamond-shaped transportation system based on the scale and intensity of population movements (Zhao et al., 2020). These four urban agglomerations (UAs) are also delineated in China's 13th Five-Year Plan (2016-2020) (Fig. 2): Beijing-Tianjin-Hebei (BTH), the Yangtze River Delta (YRD), the Pearl River

[2] China's 13th Five-Year Plan (2016-2020), 14th Five-Year Plan (2021-2025), Opinions of the CPC Central Committee and the State Council on Accelerating the Construction of a Unified National Market (2022).

[3] The pursuit of national urban agglomerations has been a central aspect of China's 'new urbanization' policy, including initiatives aimed at coordinating the development of the Beijing-Tianjin-Hebei, integrating the Yangtze River Delta, promoting growth in the Pearl River Delta, and constructing the Chengdu-Chongqing dual-city economic zone.

Delta (PRD), and Chengdu-Chongqing (CHC). These UAs comprise 64 prefectural and higher-level (direct-controlled municipalities[4] and provincial capitals) cities. These regions are of great relevance in the national economy. In 2019, they accounted for a mere 7.23% of the national land area of Mainland China, but at the same time represented 40.3% of its total population and 48% of its gross domestic product (GDP).

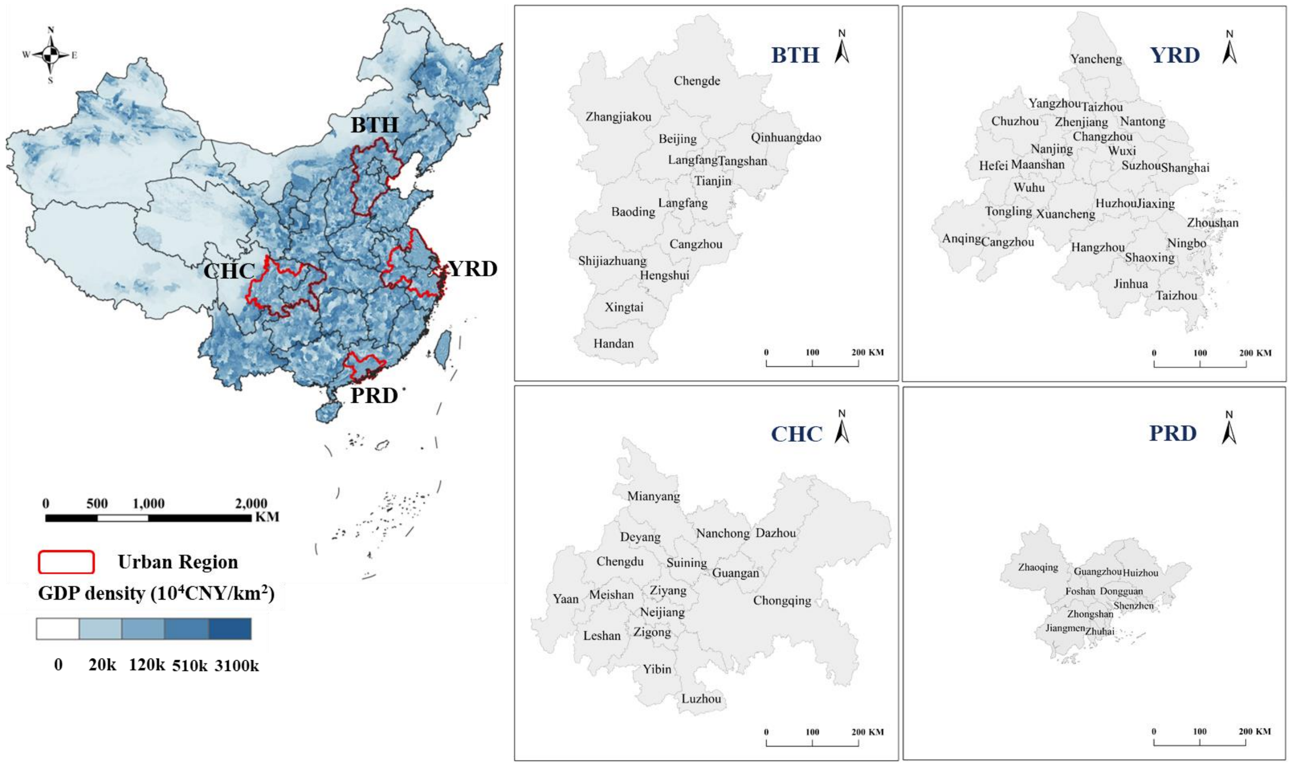

Fig. 2 Location of the four urban agglomerations in China

3.2 Data

3.2.1 Railway frequencies between cities in China's four major UAs

Due to the disruptive impact of Covid-19 on intercity transportation connections, with a 36% decrease in high-speed trains in 2020 compared to the preceding year (Huang et al., 2022), our analysis uses passenger rail frequency data for 2019 to specify inter-city transportation network connections. Since train schedules are relatively fixed, we extracted train schedules for a single day in December 2019[5]. We include both high-speed trains (designed to operate at speeds between 200 and 350 km/h) and conventional trains (designed to operate at a speed below 200 km/h). The train frequencies between cities were obtained from the *Shengming* train timetable and cross-validated with the online railway ticket system. When a city has different train stations, we merged the records. The retrieved train frequencies are constructed as four symmetric matrices across the 26, 13, 9, and 16 cities in the four UAs, respectively.

[4] The directly-controlled municipalities refer to the unique administrative divisions in China. These municipalities, including Beijing, Shanghai, Tianjin, and Chongqing, are under the direct jurisdiction of the central government and are on par with provinces in terms of administrative powers.

[5] The timing of data collection is set to filter the impact of the seasonal influences on train frequencies, include the weekends, summer/winter vacation for students and national holidays.

3.2.2 Retail price index and socio-economic indicators

RMI is measured based on the retail price index for 16 commodities in each city in 2019[6]. All price data are extracted from China's Statistical Yearbook of 2020. Since four cities' commodity price indices were unavailable, we used their consumer price indices instead. In addition, socio-economic control variables such as GDP and population are derived from the National Bureau of Statistics and China's City Statistical Yearbook of 2020.

3.3 Operationalization of variables

3.3.1 Regional market integration

Extant studies primarily draw on the price difference approach to measure market integration. This method is based on the idea that trade barriers and market fragmentation result in more significant discrepancies in commodity prices across places, in which case relative price volatility can reflect the extent of free trade activities across and the level of integration of regional markets (Parsley & Wei, 2001). Empirical studies have adopted this approach to assess the domestic and provincial market integration using commodity price differences (Ke, 2015; Su & Liang, 2021). Building on these studies, we use the price difference method to estimate RMI in the UAs. Our analysis centers on the regional markets delineated by UAs, which captures the interconnectedness of market activities and trade flows within the boundaries of UAs. We use a refined dataset of city-level commodity prices to examine how cities within regional markets function as an integrated economic entity.

The relative price differentials serve as a proxy for changes in transaction costs and market fragmentation and are derived in three consecutive steps: (1) we calculate the relative price difference of 16 commodities in each city, after which (2) we use the de-mean method to estimate the real price differentials caused by market fragmentation, to then finally (3) use standard deviations to measure the RMI of each city in its corresponding UA.

Assume that p_{it}^k, p_{jt}^k are the price of commodity *k* in city *i, j* at time *t*. Q_{ijt}^k denotes the price difference in the city pair (*i, j*) for commodity *k* at time *t* (Eq. 1). First, we calculate the first-order price difference[7] ΔQ_{ijt}^k as the real price difference in Eq. 2.

$$Q_{ijt}^k = \ln(p_{it}^k) - \ln(p_{jt}^k) \tag{1}$$

$$\left|\Delta Q_{ijt}^k\right| = \left|\ln\left(\frac{p_{it}^k}{p_{jt}^k}\right) - \ln\left(\frac{p_{it-1}^k}{p_{jt-1}^k}\right)\right| = \left|\ln\left(\frac{p_{it}^k}{p_{it-1}^k}\right) - \ln\left(\frac{p_{jt}^k}{p_{jt-1}^k}\right)\right| \tag{2}$$

Second, we calculate the real price differentials q_{ijt}^k resulting from heterogeneous market conditions[8] ε_{ijt}^k. Specifically, we employ a de-mean method to mitigate the systematic bias associated with specific product characteristics. For a given year *(t)* and a specific product category (*k*) within a region, we calculate the average price $\left|\Delta\overline{Q_t^k}\right|$ to remove the commodity difference a^k from the relative price difference $\left|\Delta Q_{ijt}^k\right|$, as expressed in Eq. 3:

$$q_{ijt}^k = \left|\Delta Q_{ijt}^k\right| - \left|\Delta\overline{Q_t^k}\right| = \left(a^k - \overline{a^k}\right) + \left(\varepsilon_{ijt}^k - \overline{\varepsilon_{ijt}^k}\right) = \left(\varepsilon_{ijt}^k - \overline{\varepsilon_{ijt}^k}\right) \tag{3}$$

And third and finally, we use the standard deviation method to calculate the market fragmentation of k types of commodities, noted as $Var(q_{ijt})$, and use this to calculate the degree of RMI as the reciprocal of the market fragmentation in Eq. 4:

---

[6] The retail price indices of 16 commodities include food, beverage and tobacco, clothing, textiles, household appliances, office articles, daily necessities, sports and recreation, transportation and communication, furniture, cosmetics, jewelry, medicine, newspaper and magazine, fuels and construction materials.

[7] The reason for calculating the first-order difference of price differentials here is twofold: 1) considering market fragmentation as a specific case of 'ice-berg' costs, the convergence of relative prices can be reflected by the first-order difference of relative prices. 2) Our original data consists of year-on-year retail price indices; the first differences can be used to construct relative price differentials that capture time-varying features.

[8] The price differences in commodities can be attributed to the heterogeneity of the products and market fragmentation. So, we need filter the unobservable effects that caused by the specific attributes of products.

$$RMI_{it} = \sqrt{\frac{1}{Var(q_{it})}} = \frac{\sum_{i \neq j}^{n-1} Var(q_{ijt})}{n-1} \qquad (4)$$

where RMI_{it} equals regional market integration in city i at time t, n refers to the total number of cities in its corresponding region, and $n-1$ represents the number of price difference pairs containing city i.

3.3.2 City network connectivity

Urban networks are inherently multiplex (Burger et al., 2014), comprising multifunctional, multidirectional, and multiscalar flows. Transportation connections are particularly important as a holistic representation of regional network connectivity: they can embody and capture the population movements and, amongst many other things, the flows of knowledge that go with these (Jiao et al., 2017). Following Zhao et al. (2022), we construct a composite transportation matrix by assigning a weight of 0.33 and 0.67 to the frequencies of conventional and high-speed trains, respectively. As degree centrality (*DC*) reflects the number of direct connections sent to/from a city (Wei et al., 2022), we measure the transportation network connections between any pair of cities i and j by *DC* as the key independent variable.

Fig. 3 visualizes the transportation network connections between 64 cities. The lines denote the train frequencies between a city pair, with warmer colors indicating more connections. The nodes represent the 64 cities in the four UAs, with node sizes varying with the degree centrality of cities (i.e., the number of trains). The figure shows that the municipalities directly under the central government (e.g., Shanghai and Chongqing), and provincial capital cities (e.g., Nanjing and Chengdu) have intensive connections within their regions. Cities in the YRD exhibit the densest connections among the four UAs.

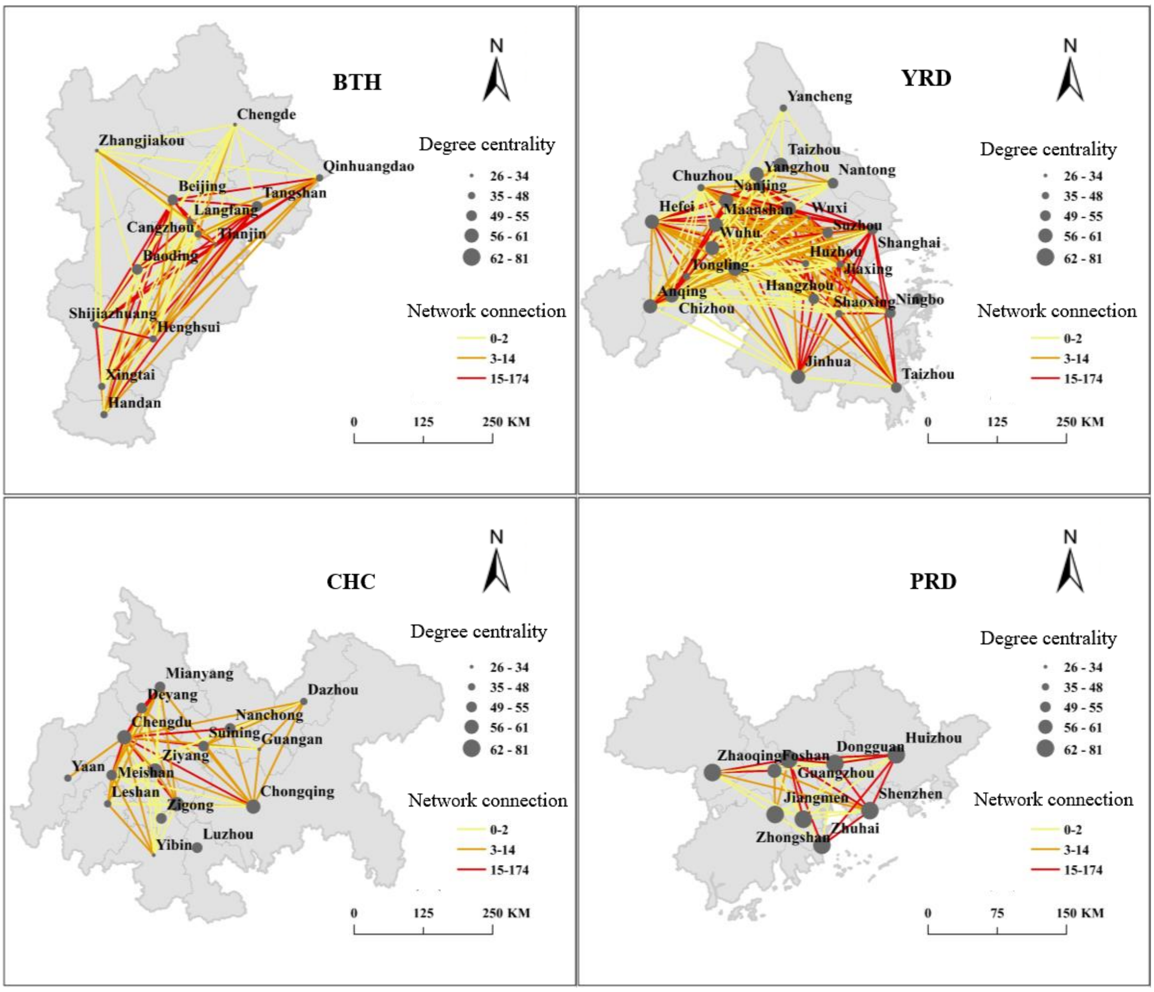

Fig. 3 Transportation network connections between cities in the four UAs

3.3.3 Measuring agglomeration externalities

Agglomeration externalities broadly comprise two different types: Marshall–Arrow–Romer (MAR) externalities and Jacobs externalities (Glaeser et al., 1992; van Meeteren et al., 2016). Empirical studies commonly use industry specialization and diversification indicators as proxies for these (see Huang et al., 2020; Cheng et al., 2022).

The crux of MAR externalities is that specialization leads to concentrations of labor, products, and information through intra-industry interactions, which can reduce transportation costs in market transaction activities (Beaudry & Schiffauerova, 2009). In this study, we employ location quotients to estimate regional industrial specialization (RSI):

$$LQ_{i,j} = \frac{E_{i,j}/E_i}{E_j/E} \tag{5}$$

$$RSI_{ij} = \max(LQ_{i,j}) \tag{6}$$

Jacobs (1969) posited that the diversity of industries in close proximity leads to increased opportunities for sharing ideas and infrastructure across industries and fosters the exchange of more diversified goods and services. The reciprocal value of the Herfindahl-Hirschman index is used to quantify regional industrial diversification (RDI):

$$RDI_{ij} = \frac{1}{\sum_j |E_{ij-}E_j|} \quad (7)$$

where $E_{i,j}$ refers to the number of employees in industry j of city i; E_i denotes the total number of employees in city i; E_j represents the total number of employees in industry j; and $LQ_{i,j}$ is the degree of industrial specification in a specialized industry j in city i.

3.3.4 Control variables

Our selection of control variables is based on the factors known to potentially influence RMI. First, economic size is measured by GDP per capita (Ke, 2015). Second, as government control can restrict regional trade activities and market integration processes (Fan et al., 2007), we include this as the share of government expenditure in GDP. Third, as cities with rich resources are less likely to engage in regional market cooperation, we introduce a proxy for resource abundance as the share of employees in the mining industry (Shao & Yang, 2010). Last, we use the share of total export and import trade in GDP to benchmark regional trade openness, as it provides access to diversified external resources (Sheng & Mao, 2011). The descriptive statistics of the variables are reported in Appendix A.

3.4 Methods

3.4.1 Benchmark model

Drawing on the study of Huang et al. (2020), we construct a benchmark model to investigate the impact of a city's transportation network connections on RMI:

$$RMI_i = \beta_0 + \beta_1 * DC_i + \beta_2 * RSI_i + \beta_3 * RDI_i + \beta_4 * X_i + \mu_i \quad (8)$$

where RMI_i is the regional market integration of city i. DC_i is the degree centrality of city i, which represents the transportation network connections of city i with other cities in its corresponding region. RSI_i and RDI_i represent the impacts of agglomeration externalities (i.e., MAR externalities and Jacobs externalities), respectively. X_i denotes a set of control variables, and μ_i is a random error term[9].

3.4.2 Spatial econometric model

Given the possible spatial interactions in the data, we conducted a spatial autocorrelation analysis on the dependent variable RMI. The global Moran's I index of RMI is 0.345 ($p<0.001$), suggesting the presence of spatial dependence in RMI. Identifying spatial effects highlights the need to include spatial models to mitigate biases in parameter estimations (Cliff & Ord, 1969). As per Anselin (1988), we start with OLS regression as the benchmark model for spatial modeling. Based on Huang et al. (2020), we estimate the spatial autoregressive model (SAR), spatial error model (SEM), and spatial Durbin model (SDM) to analyze the impact of transportation network connections on RMI. If the Lagrange Multiplier[10] (LM) and robust LM statistics are significant, the SDM is the preferred model for our analysis. The model specification is as follows.

[9] We have considered other control variables, such as openness (with FDI as indicator), infrastructure (with road area per capita as indicator) and industrial structure (with ratio of tertiary industry output/total industry output). However, the estimates of these variables are not significant. Hence, we have removed these variables from the model and manuscript.

[10] The model selection test contains two statistics: LM-Error and LM-Lag. If the LM-Error is significant, then it indicates SEM; if a significant LM-Lag statistic is observed, SAR is considered as an option. If both statistics are significant, we perform a Robust LM test to gain further insight.

$$RMI_i = \rho W_{ij} RMI_j + X_i \beta_1 + W_{ij} X_i \beta_2 + \mu_i \quad (9)$$

where RMI is the level of regional market integration; ρ and λ denote the spatially lagged coefficient and the spatial error dependence; W_{ij} is the spatial weight matrix; X_{it} is a set of explanatory variables that are hypothesized to influence RMI, as specified in Equation (8); μ_i is the random error.

3.4.3 Instruments to address endogeneity

Analyses of how cities' transportation network connections affect RMI are susceptible to potential reverse causality: integration may be a cause rather than an outcome of city linkages. For instance, cities with higher RMI generally have more contacts with other cities, which may promote resource flows and enhance cities' transportation connections in the network (Hui et al., 2011).

We address this methodological issue using instrumental variables (IVs)[11] (Meijers & Burger, 2010). A range of instrumental variables has been proposed in the literature, including lags (Su & Liang, 2021), historical data (Möller & Zierer, 2018), and transportation infrastructure (Baum-Snow et al., 2017). However, variables related to transportation are not applicable in this study: improvements in transportation infrastructure could reduce transaction costs and regional price differences. They would therefore increase cities' level of RMI.

Given that train frequencies proxy population mobility, instrumental variables should relate to population flows rather than directly affecting RMI. Drawing on Gao et al. (2019), cities with abundant tourism resources are highly connected within a region. The number of 5A scenic spots[12] can reflect a city's attractiveness (Zhou et al., 2021). Furthermore, following Ganong and Shoag (2017), the regional income disparity can drive population flows between cities. Therefore, we use the number of 5A scenic spots and historical income in 1995 as IVs[13].

4 Empirical results

4.1 Exploring the presence of an inverted 'U'-shaped effect

We commence with a discussion of our benchmark models (Table 1). In these models, the signs of all control variables are as expected across all specifications. Agglomeration externalities (*RSI* and *RDI*) and international openness (*OPEN*) are positively associated with cities' level of market integration. The negative coefficients of *GDP, GOV,* and *RES* suggest that economic development, government control, and resource endowment negatively impact cities' RMI, echoing findings in earlier studies (Fan et al., 2007; Ke, 2015). Based on the residuals of the OLS model (Model 1), we employ SAR, SEM, and SDM for the spatial effects estimation (Models 2-3-4). The Lagrange multiplier and Robust Lagrange multiplier tests are significant at 1% levels, suggesting that SDM (Model 4) should be selected.

Table 1 Results of the basic estimations

Variables	Model (1)	Model (2)	Model (3)	Model (4)
DC (log)	12.28(4.29)***	2.13(2.90)	3.92(4.57)	8.59(3.62)**
DC^2 (log)	-1.73(0.50)***	-0.41(0.45)	-0.81(0.59)	-0.99(0.50)**
RSI	2.06(0.64)***	1.93(0.68)***	1.89(0.71)***	1.97(0.82)***

[11] A valid instrumental variable should be correlated to the endogenous variables and not to the dependent variable.

[12] Under the rating system for tourist attractions in China, 5A (AAAAA) scenic spots represent the highest level of tourist scenic spots certificated by the National Tourism Administration of China.

[13] The "Labor Law" that took effect in China on January 1, 1995 established a minimum wage system and set corresponding wage standards, which have had a significant impact on the regional labor mobility. Thus, we select the average level of income in 1995.

RDI	2.56(1.07)**	1.46(0.78)*	2.25(1.21)*	1.08(0.59)*
GDP (log)	-11.59(4.21)**	-11.83 (4.47)***	-11.98(4.39)***	-12.36(4.20)***
GOV	-1.38(0.30)***	-1.04(0.29)***	-1.35(0.30)***	-0.87(0.30)***
RES	-8.17(2.65)***	-4.33(2.23)*	-5.33(3.19)*	-3.87(1.98)*
OPEN	0.14(0.04)***	0.10(0.05)**	0.15(0.04)***	0.03(0.02)*
W*RMI				0.28(0.45)
W*DC (log)				33.24(15.21)**
W*DC^2 (log)				-3.5(1.76)**
W*RSI				1.13(1.54)
W*RDI				0.66(3.08)
W*GDP (log)				-7.23(3.33)**
W*GOV				-2.78(0.76)***
W*RES				-3.39(10.54)
W*OPEN				0.16(0.14)
Constant	150.95(47.22)***	180.10(47.66)***	173.56(47.45)***	194.11(45.06)***
Adj. R^2	0.53			
F-statistic	10.06			
ρ/λ		0.68(0.23)***	0.65(0.26)***	0.87(0.23)***
Log-L		-220.33	-222.54	-203.16
Obs.	64	64	64	64

Notes: 1) Significance: 0.01 (***), 0.05 (**), and 0.1 (*).

2) Robust standard errors of parameter estimates are reported in parentheses.

As put forward in Hypothesis 1, the inverted U-shaped relationship between transportation network connections and RMI is supported by Model 4. To confirm the validity of this finding, we employed the UTEST estimation[14] developed by Lind and Meholum (2010). The results, reported in Appendix B, support the presence of an inverted U-shaped relationship. The significantly positive linear and negative quadratic terms of transportation network connections suggest that cities with more connections experience increased RMI. Still, beyond a certain threshold, the adverse effects of congestion and competition begin to outweigh the positive impact of access to markets and resources. This finding is consistent with previous research on the diminishing returns of network connections (Cao et al., 2022).

4.2 Indirect effect: Spatial spillovers of transportation network connections

Hypothesis 2 proposes that a city's transportation network connections are expected to generate spatial spillover effects on connected cities. The results of Model 4 provide support for this hypothesis, as the spatial lag coefficient of ρ and the significant *W*DC* and *W*DC^2* terms indicate that the transportation network connections of a city have a spatial spillover effect on the linked city. This is consistent with the findings of Huang et al. (2020), who observed that urban transportation network connections could generate cross-spatial spillovers. Specifically, in the initial phase, where transportation network connections positively impact RMI, RMI's performance can be enhanced in connected cities by providing access to additional resources and markets. Put differently: cities are 'borrowing size' from other cities by being well embedded in regional networks. This was also empirically shown

[14] The UTEST gives an exact test of the presence of an inverted U-shaped relationship between a predictor and a response variable on a specific interval after the benchmark estimation model.

in Meijers et al. (2016), who examined cities' transportation networks in Western Europe. However, when transportation network connections lead to negative competition and congestion effects, there can be negative spillovers on connected cities, i.e., the dominance of 'agglomeration shadows.'

4.3 Heterogeneity: The importance of size for different cities

Following Meijers et al. (2016), we use population size to proxy the size of cities. To determine whether the impact of transportation network connections varies across cities with different sizes, we introduce the interaction terms *size*DC* and *size*DC²* (Table 2). To validate the heterogenous impact of different city sizes, we adopt the method of Meijers et al. (2016), dividing city size into four quartiles (Appendix C). The coefficients of *size*DC (negative)*, *size* DC² (positive)* in model 4 indicate that, as city size grows, the positive impact of transportation network connections decreases, while the negative effect is amplified. Results of different city size quantiles show that an increase in city size from the 25th to the 75th percentile led to a decrease in the positive effect and an increase in the negative effect on RMI. Hypothesis 3 is, therefore, accepted.

Specifically, when transportation network connections positively influence RMI, smaller cities experience more substantial positive effects than larger ones. In other words, smaller cities are more likely to 'borrow size' and reap the benefits through positive network spillovers (Meijers et al., 2016). This finding aligns with the study of Camagni et al. (2016), who argued that smaller cities derive more advantages from networking than larger cities. As transportation network connections reach a threshold, the concentration of labor and capital gives rise to a range of negative externalities in larger cities, so they are more prone to experiencing 'agglomeration shadows,' i.e., strong negative effects on their RMI.

Table 2 Estimates of the regional heterogeneity

Variables	Model (1)	Model (2)	Model (3)	Model (4)
DC (log)	14.01(4.79)***	10.42(4.74)**	15.15(3.82)***	7.11(3.74)*
DC^2 (log)	-1.96(0.57)***	-1.34(0.59)**	-1.80(0.47)***	-0.89 (0.46)*
RSI	2.01(0.70)***	1.73(0.74)**	1.46(0.60)**	1.43(0.76)*
RDI	2.51(1.06)**	1.34(0.69)*	1.31(1.02)	0.84(0.49)*
GDP (log)	-11.32(4.14)***	-10.26 (4.09)**	-7.64(3.33)**	-16.53(4.51)***
GOV	-1.41(0.29)***	-1.16(0.29)***	-0.99(0.25)***	-1.18(0.32)***
RES	-8.06(2.59)***	-5.56(3.14)*	-6.19(2.47)**	-4.33 (2.60)*
OPEN	0.14(0.04)***	0.10(0.05)**	0.06(0.03)*	0.03(0.02)*
Size*DC	-0.75(0.28)**	-0.55(0.03)*	-0.43(0.24)*	-0.15 (0.07)**
Size* DC^2	0.11(0.04)***	0.04(0.02)*	0.07(0.04)*	0.02(0.01)*
W*RMI				0.09(0.50)
W*DC (log)				29.01(9.97)***
W*DC^2 (log)				-8.95(2.35)***
W*RSI				1.27(2.03)
W*RDI				3.98(4.20)
W*GDP (log)				7.76(18.69)
W*GOV				-1.19(1.55)
W*RES				-13.42(11.51)
W*OPEN				0.13(0.17)

W*(Size*DC)				-2.11(1.10)*
W*(Size* DC^2)				0.32(0.19)*
Constant	146.68(45.78)***	119.97(47.62)**	83.86(40.72)**	52.63(15.96)***
Adj. R^2	0.54			
F-statistic	8.93			
ρ/λ		0.54(0.19)***	0.81(0.12)***	0.77(0.23)***
Log-L		-212.69	-206.19	-191.58
Obs.	64	64	64	64

Notes: 1) Significance: 0.01 (***), 0.05 (**), and 0.1 (*).

2) Robust standard errors of parameter estimates are reported in parentheses.

4.4 Investigating the (possible) interaction effect of network and agglomeration externalities

Hypothesis 3 suggests the presence of either a complementary or a substitution effect between network externalities and agglomeration externalities on RMI. In Table 3 (Model 1-4), the interaction terms (*RSI*DC, RSI* DC^2, RDI*DC,* and *RDI** DC^2) are not significant: our estimates do not provide evidence of the presence of complementary or substitution effects of network externalities for agglomeration externalities. Hypothesis 3 is therefore rejected. This finding is consistent with the study by Huang et al. (2020), arguing that there are no interactions between network externalities and agglomeration externalities for Chinese cities' performance.

Table 3 Estimates of interactions between network externalities and agglomeration externalities

Variables	Model (1)	Model (2)	Model (3)	Model (4)
DC (log)	13.31(5.78)**	1.05(3.23)	2.17 (5.17)	8.20(4.26)*
DC^2 (log)	-1.88(0.68)***	-0.23(0.51)	-0.58(0.68)	-0.98(0.56)*
RSI	2.40(0.88)***	2.12(1.07)**	2.48(1.05)**	1.58(0.95)*
RDI	2.07(1.19)*	1.41(0.8)*	2.16(1.27)*	1.14(1.07)
GDP (log)	-10.50(4.34)**	-10.52 (4.81)**	-11.05(4.37)**	-13.94(4.14)***
GOV	-1.32(0.32)***	-0.95(0.30)***	-1.24(0.31)***	-0.84(0.32)***
RES	-8.24(3.00)***	-5.03(3.37)	-7.50(3.53)**	-3.87(2.51)*
OPEN	0.14(0.06)**	0.07(0.06)	0.13(0.06)**	0.03(0.02)*
RSI*DC	-0.01(0.01)	-0.04(0.07)	-0.01(0.01)	0.01(0.01)
RSI* DC^2	6.54e-06(5.76e-06)	2.96e-06(7.02e-06)	7.97e-06(7.47e-06)	4.71e-06(6.21e-06)
RDI*DC	0.01(0.01)	-0.01(0.01)	-0.01(0.02)	0.01(0.01)
RDI* DC^2	-9.56e-07(0.0001)	9.30e-06(0.0001)	2.74e-06(0.0001)	0.0001(0.0001)
W*RMI				0.25(0.43)
W*DC (log)				62.76(21.79)***
W*DC^2 (log)				-6.93(2.51)***
W*RSI				5.84(3.05)*
W*RDI				1.79(3.71)
W*GDP (log)				-10.35(5.65)*
W*GOV				-2.59(0.83)***
W*RES				-8.41(11.71)
W*OPEN				0.17(0.15)
W*(RSI*DC)				0.03(0.02)

W*(RSI* DC^2)				0.01(0.01)
W*(RDI*DC)				0.03(0.04)
W*(RDI* DC^2)				0.01(0.01)
Constant	142.58(48.59)***	166.30(49.84)***	166.31(47.21)***	211.46(44.07)***
Adj. R^2	0.52			
F-statistic	8.18			
ρ/λ		0.68(0.22)***	0.59(0.17)***	1.00(0.28)***
Log-L		-219.20	-221.36	-197.67
Obs.	64	64	64	64

Notes: 1) Significance: 0.01 (***), 0.05 (**), and 0.1 (*).

2) Robust standard errors of parameter estimates are reported in parentheses.

4.4 Endogenous and robustness test

To ensure that the direction of causality runs from cities' transportation network connections to RMI, we introduced the instrumental variable approach and the spatial autoregressive model with spatial autoregressive disturbances and IVs to address the endogeneity issue. To examine the robustness of our empirical results, four different approaches are used: 1) regression using the random forest method; 2) proxying transportation network connections with high-speed rail frequencies; 3) estimating with generalized spatial two-stage least square techniques; 4) introducing the spatial inverse distance matrix as an alternative. Our findings stand robust across different estimations, as evidenced by the detailed descriptions provided in Appendices D and E.

5 Discussion and conclusions

Urban networks have increasingly attracted interest in geography and regional science. They are often used to provide external benefits that transcend geographic proximity. Although earlier research has examined the impact of urban networks on cities' economic performance, there has been limited research on their effect on RMI. This paper presented such an analysis, enhancing our understanding of how network connections influence RMI in the context of China's leading UAs.

We commenced by constructing a conceptual framework linking transportation network connections and RMI through the channels of 'sharing' and 'matching,' proposing four research hypotheses and scrutinizing them through explicitly spatial models. Our main findings are that:

1. A city's transportation network connections have an inverted 'U'-shaped effect on its level of RMI. Cities with more connections tend to experience increased RMI, but once they meet a certain threshold, the negative effects of congestion and competition begin to outweigh the positive impact. Our finding aligns with previous research on the diminishing returns of network connections (Cao et al., 2022) and extends this insight by relating transportation network connections to regional markets.
2. A spatial Durbin model captured urban network externalities in China's leading UAs, suggesting that transportation network connections can generate spatial spillover effects. Our results show that the benefits of transportation network connections for RMI are not geographically constrained but can be shared through networking.
3. The impact of cities' transportation network connections on RMI varies by city size. Smaller cities are more likely to borrow benefits from another market through transportation networks, while larger cities tend to be dominated by agglomeration shadow effects. The literature on the relative benefits of urban networks for larger and smaller cities is diverse and inconclusive. Studies such as Camagni et al. (2016) posited that smaller

cities might derive more significant productivity advantages from cooperative relationships than their larger counterparts. Meijers et al. (2016) employed alternative measures of networks and found that smaller cities may be less capable of attaining urban functions through these networks. These different findings underscore the necessity for further research to account for both the multi-faceted nature of networks (multiplexity) and the diverse characteristics of cities of varying sizes (regional heterogeneity).

4. Neither complementary nor substitution effects between network externalities and agglomeration externalities were observed. This is consistent with the findings of Huang et al. (2020), arguing that there is no complementary or substitution effect between network and agglomeration externalities for Chinese cities' performance. However, using data on knowledge collaborations, Yao and Song (2019) found evidence of both complementary and substitution effects in Chinese cities. This discrepancy in findings may be attributed to the multiplexity of network connections. The geographic scope of transportation network connections, proxied by train frequencies, is typically more circumscribed than knowledge collaboration between regions, potentially influencing our findings.

The results have implications for understanding RMI from the perspective of urban networks. We show how smaller cities can gain critical urban mass by being well-connected in regional networks. Furthermore, our findings point to the importance of one of the driving forces of RMI. As a product of territorially embedded spaces, inter-city transportation network connections are essential to reduce trade frictions associated with long distances between cities. For example, improvements in transportation infrastructure networks generally lead to reduced trade costs: it can facilitate the mobility of goods, people, capital, and information and unleash the market potential within cities/regions. However, our finding of the inverted-U-shaped relationship between transportation network connections and RMI also highlights the need to balance the benefits of the network and the costs of congestion.

Translating these findings into policy recommendations, it can be argued that given geographical restrictions on factor flows and the regional market fragmentation in China, actions aimed at fostering RMI through intercity transportation network connections should be implemented. First, policy endeavors should be directed towards promoting regional integration through infrastructure networks, such as increasing infrastructure investments to develop an integrated regional transportation network anchored in cities for regional trade activities. Second, as Capello (2000) suggested, reaping benefits from network externalities also requires proactive network participation. This means policymakers, especially in smaller cities, should be open to embracing the organizational changes necessary to achieve network externalities. For example, smaller cities can collaborate with other cities through networks to leverage their specialized functions and undertake production industries that have been shifted from larger cities. While larger cities can work on proactive policies to gain market information, foster innovation (Puga, 2010), and leverage size-related advantages of (re)combining abundant capital and resources through networks (Cao et al., 2022). This could involve attracting investment in trading centers and technology hubs and expanding access to a diverse pool of specialties and professionals. Third, particularly important for policy practices, is the 'integrated' governance frames. The normative agendas proposed by central and local governments concerning integration are sometimes argued to be a mere 'blueprint' (Luo & Shen, 2008, p. 211), with a lack of power to implement the planning policy directly (Harrison & Gu, 2021). It is therefore essential to translate policy objectives into concrete governance measures, such as the establishment of 'The Yangtze River Delta Pilot Free Trade Zone Alliance' and 'integrated transportation system,' echoing the practice of promoting regional free trade and reducing trade friction.

Future research can build on and enhance our research in several ways. First, the cross-sectional approach of this study has the advantage of identifying the general impacts of cities' transportation network connections. Our conceptual framework of network connectivity and RMI opens up new opportunities for panel studies investigating the time-varying role of urban networks. Second, another extension could explore alternative instrumental variables to address endogeneity concerns in urban network studies. The instruments used in our analysis are relevant to the network indicators under investigation, and they were selected based on the available data and factors affecting intercity population flows. However, one can resort to other variables (e.g., historical network indicators), which may help further elucidate the rigor of the causal relationship. Third, our study focuses exclusively on the impact of intra-regional transportation network connections. Subsequent studies could incorporate the potential effects of inter-regional transportation network connections on regional market activities. Moreover, we observed the presence of network externalities through spatial spillover effects, while future research could explore an alternative application of methods to operationalize network externalities. Lastly, the complexity and multiplexity of regional networks should be given due attention. Future studies can investigate a much wider variety of urban networks beyond transportation connections, and promising lines of inquiry could examine the heterogenous impact of different types of urban networks, such as corporate, information, migration, and innovation networks, on the broad socio-economic performance within the region at large.

Reference

Alonso, W. (2013). Urban zero population growth. Daedalus. 102: 191-206.

Amin, A., & Thrift, N. (1992). Neo-Marshallian Nodes in Global Networks. *International Journal of Urban and Regional Research*, 16(4), 571-587.

Anselin, L. (1988). Spatial econometrics: methods and models (Vol. 4). Springer Science & Business Media.

Bathelt, H., Malmberg, A., & Maskell, P. (2004). Clusters and knowledge: local buzz, global pipelines and the process of knowledge creation. *Progress in Human Geography*, *28*(1), 31-56.

Batten, D. F. (1995). Network cities: creative urban agglomerations for the 21st century. *Urban Studies*, 32(2), 313-327.

Baum-Snow, N., & Turner, M. A. (2017). Transport Infrastructure and the Decentralization of Cities in the People’s Republic of China. *Asian Development Review*, 34(2), 25-50.

Beaudry, C., & Schiffauerova, A. (2009). Who’s right, Marshall or Jacobs? The localization versus urbanization debate. *Research Policy*, 38(2), 318-337.

Burger, M. J., Van Der Knaap, B., & Wall, R. S. (2014). Polycentricity and the multiplexity of urban networks. *European Planning Studies*, 22(4), 816-840.

Burger, M. J., Meijers, E. J., Hoogerbrugge, M. M., & Tresserra, J. M. (2015). Borrowed size, agglomeration shadows and cultural amenities in North-West Europe. *European Planning Studies*, *23*(6), 1090-1109.

Burger, M. J., & Meijers, E. J. (2016). Agglomerations and the rise of urban network externalities. *Papers in Regional Science*, *95*(1), 5-15.

Camagni, R., Capello, R., & Caragliu, A. A. (2016). Static vs. dynamic agglomeration economies. Spatial context and structural evolution behind urban growth. *Papers in Regional Science*, 95(1), 133-159.

Capello, R., & Camagni, R. (2000). Beyond optimal city size: an evaluation of alternative urban growth patterns. *Urban Studies*, 37(9), 1479-1496.

Capello, R. (2000). The city network paradigm: measuring urban network externalities. *Urban Studies*, *37*(11), 1925-1945.

Capello, R. (2009). Spatial spillovers and regional growth: a cognitive approach. *European Planning Studies*, 17(5), 639-658.

Cao, Z., Derudder, B., Dai, L., & Peng, Z. (2022). ‘Buzz-and-pipeline’ dynamics in Chinese science: the impact of interurban collaboration linkages on cities’ innovation capacity. *Regional Studies*, *56*(2), 290-306.

Castells, M. (2000). Toward a sociology of the network society. *Contemporary Sociology*, *29*(5), 693-699.

Cheng, Y., & LeGates, R. (2018). China's hybrid global city region pathway: Evidence from the Yangtze River Delta. *Cities*, 77, 81-91.

Cheng, Z., Li, X., Zhu, Y., & Wang, M. (2022). The effects of agglomeration externalities on urban green total-factor productivity in China. *Economic Systems*, 101025.

Cliff A.D., & Ord J.K. (1969). The problem of spatial autocorrelation. A.J. Scott (Ed.), London Papers in Regional Science, *Studies in Regional Science*, 1, Pion, London, 25-55.

Dawson, P. J., & Dey, P. K. (2002). Testing for the law of one price: Rice market integration in Bangladesh. *Journal of International Development*, 14(4), 473-484.

Derudder, B., Meijers, E., Harrison, J., Hoyler, M., & Liu, X. (2022). Polycentric urban regions: conceptualization, identification and implications. *Regional Studies*, 56(1), 1-6.

Duranton, G., & Puga, D. (2004). Micro-foundations of urban agglomeration economies. In *Handbook of Regional and Urban Economics*, 4, 2063-2117.

Fan, A. J., Li, Z., Liu X. Y. (2007). The empirical study on market segmentation of China and its influence factors based on commodity market. *Nankai Economic Studies.* (5). 111-119.

Fang, C., & Yu, D. (2017). Urban agglomeration: An evolving concept of an emerging phenomenon. *Landscape and Urban Planning*, 162, 126-136.

Fossati, S., Lorenzo, F., & Rodriguez, C. M. (2007). Regional and international market integration of a small open economy. *Journal of Applied Economics*, 10(1), 77-98.

Ganong, P., & Shoag, D. (2017). Why has regional income convergence in the US declined? *Journal of Urban Economics*, 102, 76-90.

Gao, Y., Su, W., & Wang, K. (2019). Does high-speed rail boost tourism growth? New evidence from China. *Tourism Management*, 72, 220-231.

Glaeser, E. L., Kallal, H. D., Scheinkman, J. A., & Shleifer, A. (1992). Growth in cities. *Journal of Political Economy*, 100(6), 1126–1152.

Harrison, J., & Gu, H. (2021). Planning megaregional futures: Spatial imaginaries and megaregion formation in China. *Regional Studies*, 55(1), 77-89.

Herranz-Loncán, A. (2007). Infrastructure investment and Spanish economic growth, 1850-1935. *Explorations in Economic History,* 44(3), 452-468.

Huang, Y., Hong, T., & Ma, T. (2020). Urban network externalities, agglomeration economies and urban economic growth. *Cities*, *107*, 102882.

Huang, J., Wang, J., & Li, Y. (2022). How Covid-19 affects intercity transportation in China: high-speed rail versus air. *Regional Studies*, *Regional Science*, 9(1), 409-411.

Hui, E. C. M., Wong, F. K. W., Li, S. M., & Yu, K. H. (2011). Integrations, identity and conflicts: A cross-border perspective on residential relocation of Hong Kong citizens to Mainland China. *Habitat International*, 35(1), 74-83.

Jacks, D. S. (2005). Intra-and international commodity market integration in the Atlantic economy, 1800–1913. *Explorations in Economic History*, 42(3), 381-413.

Jacobs, J. (1969). The economy of cities. Vintage Books: A Division of Random House.

Jiao, J., Wang, J., & Jin, F. (2017). Impacts of high-speed rail lines on the city network in China. *Journal of Transport Geography*, 60, 257-266.

Johansson, B., & Quigley, J. M. (2004). Agglomeration and networks in spatial economies. In Fifty years of regional science. Springer, Berlin, Heidelberg. 165-176.

Johansson, B. (2005). Parsing the menagerie of agglomeration and network externalities. Industrial clusters and inter-firm networks. *Edward Elgar, Cheltenham,* 107-147.

Ke, S. (2015). Domestic market integration and regional economic growth-China's recent experience from 1995–2011. *World Development*, *66*, 588-597.

Liu, X., Derudder, B., & Wu, K. (2016). Measuring polycentric urban development in China: An intercity transportation network perspective. *Regional Studies*, 50(8), 1302-1315.

Lu, J., & Mao, W. F. (2020). The Rise of Urban Network Externalities: A New Mechanism for the High-quality Integrated Development of Regional Economy. *Economists*, 12, 62-70.

Luo, X., & Shen, J. (2008). Why city-region planning does not work well in China: The case of Suzhou–Wuxi–Changzhou. *Cities*, 25(4), 207-217.

Mačiulis, A., Vasiliauskas, A. V., & Jakubauskas, G. (2009). The impact of transport on the competitiveness of national economy. *Transport*, 24(2), 93-99.

Meijers, E. J., & Burger, M. J. (2010). Spatial structure and productivity in US metropolitan areas. *Environment and Planning A*, 42(6), 1383-1402.

Meijers, E. J., Burger, M. J., & Hoogerbrugge, M. M. (2016). Borrowing size in networks of cities: City size, network connectivity and metropolitan functions in Europe. *Papers in Regional Science*, *95*(1), 181-198.

Möller, J., & Zierer, M. (2018). Autobahns and jobs: A regional study using historical instrumental variables. *Journal of Urban Economics*, 103, 18-33.

Niu, D., Sun, W., & Zheng, S. (2020). Travel costs, trade, and market segmentation: Evidence from China's high-speed railway. *Papers in Regional Science*, 99(6), 1799-1825.

Partridge, M. D., Rickman, D. S., Ali, K. & Olfert, M. R. (2009). Do new economic geography agglomeration shadows underlie current population dynamics across the urban hierarchy? *Papers in Regional Science*, 88(2), pp. 445–466.

Parsley, D. C., & Wei, S. J. (2001). Explaining the border effect: the role of exchange rate variability, shipping costs, and geography. *Journal of International Economics*, 55(1), 87-105.

Poncet, S. (2005). A fragmented China: Measure and determinants of Chinese domestic market disintegration. *Review of International Economics*, 13(3), 409-430.

Puga, D. (2010). The magnitude and causes of agglomeration economies. *Journal of regional science*, 50(1), 203-219.

Que, W., Zhang, Y., Liu, S., & Yang, C. (2018). The spatial effect of fiscal decentralization and factor market segmentation on environmental pollution. *Journal of Cleaner Production,* 184, 402-413.

Samuelson, P. A. (1952). Spatial price equilibrium and linear programming. *The American Economic Review*, 42(3), 283-303.

Shao, S., & Yang, L. (2010). Natural resource abundance, resource industry dependence and regional economic growth in China. *Management World*, (9), 26-44.

Sheng B., & Mao Q. L. (2011). Domestic Market Integration and Interprovincial Economic Growth in China: 1985~2008. *The Journal of World Economy,* (11): 44-66.

Shi, S., Wall, R., & Pain, K. (2019). Exploring the significance of domestic investment for foreign direct investment in China: A city-network approach. *Urban Studies*, *56*(12), 2447-2464.

Su, H., & Liang, B. (2021). The impact of regional market integration and economic opening up on environmental total factor energy productivity in Chinese provinces. *Energy Policy*, 148, 111943.

Tian, M., Wang, Y., & Wang, Y. (2022). High-speed rail network and urban agglomeration economies: Research from the perspective of urban network externalities. *Socio-Economic Planning Sciences,* 101442.

Van Meeteren, M., Neal, Z., & Derudder, B. (2016). Disentangling agglomeration and network externalities: A conceptual typology. *Papers in Regional Science*, 95(1), 61-80.

Volgmann, K., & Rusche, K. (2020). The geography of borrowing size: Exploring spatial distributions for German

urban regions. *Tijdschrift voor Economische en Sociale Geografie*, 111(1), 60-79.
Wang, Z., & Wei, W. (2022). Principles and Influential Factors of Cross-Border Economic Cooperation Zones. In Cross-Border Economic Cooperation Between China and Southeast Asian Countries. Springer, Singapore, 165-181.
Wei, Y., Wang, J., Zhang, S., Sono, D., & Xiu, C. (2022). Urban positionality in the regional urban network: Through the lens of alter-based centrality and national-local perspectives. *Habitat International*, *126*, 102617.
Wu, F. (2016). China's emergent city-region governance: a new form of state spatial selectivity through state-orchestrated rescaling. *International Journal of Urban and Regional Research*, 40(6), 1134-1151.
Yao C. C., & Song D. L. (2019). Borrowed-size, network externalities and agglomeration economies in the urban agglomerations. *Industrial Economics Research*. 2, 76-78.
Young, A. (2000). The razor's edge: Distortions and incremental reform in the People's Republic of China. *The Quarterly Journal of Economics*, 115(4), 1091-1135.
Zhao J. L., Zhang X. B., & Ren J. M. (2022). Spatial structure and influencing factors of urban network in the Yellow River Basin based on multiple flows. *Scientia Geographica Sinica*, 42(10):1778-1787.
Zhou, B., Wen, Z., & Yang, Y. (2021). Agglomerating or dispersing? Spatial effects of high-speed trains on regional tourism economies. *Tourism Management*, 87, 104392.

Appendix

Appendix A Descriptive statistics for variables

Table 1 Summary statistics of variables

Variable	Definition	Obs	Mean	Std. dev.	Min.	Max.
RMI	Regional market integration	64	53.06	10.83	26.52	80.40
DC	Degree centrality	64	215.16	259.55	0	1141.76
RSI	Industrial specialization	64	1.75	1.94	0.00067	11.52
RDI	Industrial diversification	64	0.013	0.0013	0.0089	0.016
GDP	GDP per capita	64	87439.13	46527.41	28707	203489
GOV	Government control	64	17.15	5.66	8.76	39.40
RES	Resource endowment	64	0.16	0.33	0	1.78
OPEN	Trade openness	64	30.51	34.05	.00026	145.56

Appendix B Results of UTEST estimation

We conduct a UTEST estimation based on the results of the benchmark OLS model. The results in Table 2 reveal that the extreme point calculated is 3.55, and the value range of *logDC* is [1.4563, 7.0403]. The value of the extreme point falls within the data value interval and can reject the null hypothesis at the 1% statistical level. The slope of the upper bound has a negative value in the interval, indicating that transportation network connections and RMI have an inverted U-shaped relationship.

Table 2 Results of UTEST estimation

Extreme/Turning point	3.550081	
	Lower bound	Upper bound
Interval	1.4563	7.0403
Slope	7.2449	-12.0768
t-value	2.4815	-3.8980
P>t	0.0081	0.0001
t-value	2.48	
P>t	0.00814	

Test: H1: Inverse U shape vs. H0: Monotone or U shape.

Appendix C Results of the regional heterogeneity

Table 3 Estimations of regional heterogeneity: different quantiles of city sizes

Variables	25th percentile	50th percentile	75th percentile
DC(log)	16.56(8.03)*	12.65(5.16)**	7.98(3.04)**
DC^2(log)	-1.79(0.90)*	-2.97(1.39)**	-3.31(1.57)**
RSI	2.25(3.93)	5.37(2.23)**	6.27(2.78)**
RDI	2.13(3.95)	4.80(3.86)	3.55(1.88)*
GDP(log)	-5.22 (2.54)*	-6.18(2.64)**	-4.43(2.07)*
GOV	-1.27(0.51)*	-0.83(0.46)*	-1.59(0.71)*
RES	-0.01(2. 16)	-2.32(1.07)*	-3.17(2.65)
OPEN	0.17 (0.08)*	0.09(0.12)	0.17(0.25)
Constant	71.92 (91.37)	102.85(192.36)	166.23(175.68)

Notes: 1) Significance: 0.01 (***), 0.05 (**), and 0.1 (*).
2) Standard errors of parameter estimates are reported in parentheses.

Appendix D Dealing with endogeneity

The tests of our IV regressions are reported in Table 4, alongside the standard tests for the variables' relevance and exogeneity. The weak identification test of the Cragg-Donald Wald F statistic is 13.44 for 5A scenic spots and 18.47 for historical regional income, which are greater than the critical values of Stock-Yogo at a 10% maximal IV size at a 5% confidence interval. The results suggest the IVs are effective. From the endogeneity test results, both statistics are not significant at a 10% confidence level, so it can be concluded that our key variable of interest – *DC* – can be treated as exogenous.

Table 4 Instrumental variable tests

	Result
Weak instrument test	
Stock-Yogo test/F-statistics	

	Result
Scenic spots	13.44
Historical income	18.47
Overidentification test	
Sargan's test	0.7969
Endogeneity test	
Durbin-Wu-Hausman chi-sq.	0.1123
Wu-hausman	0.1268

We reported the second-stage result of IV regression in Table 5 (see Model 1). Following Pan et al. (2020), we also apply the IVs in the SARAR model (see Model 2-4); our results provide consistent and compelling evidence for the main findings and address the spatial autocorrelation and endogeneity issues.

Table 5 Results of the SARAR-IV model

Variables	Model (1)	Model (2)	Model (3)	Model (4)
DC(log)	19.94(10.11)*	12.24(6.11)**	16.75(8.62)*	12.62(5.04)**
DC^2(log)	-2.27(1.21)*	-1.23(0.69)*	-1.90(1.00)*	-1.33(0.60)**
RSI	2.53(1.05)**	1.25(0.55)**	1.33(0.45)***	1.27(0.58)**
RDI	0.32(0.17)*	0.01(1.16)	0.42(0.91)	0.55(0.88)
GDP(log)	-6.37(1.25)***	-4.22(1.34)***	-5.93(2.59)**	-4.43(2.44)*
GOV	-1.19(0.59)**	-0.71(0.26)***	-0.84(0.20)***	-0.64(0.21)***
RES	-9.96(5.11)	-4.08(2.01)**	-4.60(2.37)*	-3.17 (1.65)*
OPEN	0.01(0.09)	0.01(0.04)	0.01(0.04)	0.01(0.03)
Size*DC			-0.61(0.23)***	
Size*DC^2			0.10(0.04)***	
DC*RSI				-0.01(0.04)
DC*RDI				1.11e-06(4.70e-06)
DC^2*RSI				-0.01(0.1)
DC^2*RDI				4.07e-06(0.00001)
Constant		24.09 (12.15)*	37.25(19.38)**	27.08(19.26)*
λ		0.87(0.17)***	0.74(0.18)***	0.86(0.16)***
ρ		0.78(0.27)***	0.70(0.25)***	0.81(0.17)***
Obs.		64	64	64

Notes: 1) Significance: 0.01 (***), 0.05 (**), and 0.1 (*).

2) Standard errors of parameter estimates are reported in parentheses.

Appendix E Robustness tests

We use four different methods to examine the robustness of our empirical results. First, as the random forest method outperforms the traditional multivariate linear regression in fitting and prediction (Cootes et al., 2012), we adopt the random forest method to randomly fit 100 rounds to generate a predicted dependent variable RMI and conduct the regression analysis. Second, as China's high-speed rail enhances regional connectivity, we use different independent variables, i.e., the degree centrality of urban network constructed by high-speed rail frequencies to proxy transportation network connections (Huang et al., 2020). Third, we use the spatial estimation method of generalized spatial two-stage least squares, which makes the spatial inference more reliable (Kelejian & Prucha, 1998; Jin & Lee, 2013). And finally, we introduced the inverse distance matrix as an alternative. Detailed results are tabulated in Table 6. The changing directions of coefficients are consistent with our estimates, indicating that the main findings also remain robust with different estimations.

Table 6 Estimates of robustness tests

Variables	Model (1)	Model (2)	Model (3)	Model (4)
DC (log)	8.07(2.78)***	7.73(3.21)**	9.24(3.78)***	3.72(2.1)*
DC^2 (log)	-1.12(0.32)***	-1.16(0.63)*	-1.19(0.50)***	-0.60(0.31)*
RSI	1.51(0.42)***	1.88(0.70)***	2.24(0.68)***	1.73(0.76)**
RDI	1.60(0.71)**	2.24(1.05)**	1.64(1.09)	0.94(1.04)
GDP (log)	-7.21(2.65)***	-10.32(4.88)**	-13.28(4.14)***	-10.73(4.85)**
GOV	-1.02(0.18)***	-1.23(0.43)***	-1.09(0.28)***	-0.40(0.21)*
RES	-5.87(1.83)***	-7.79(2.65)***	-4.32(2.24)*	-4.70(3.85)
OPEN	-0.10(0.03)***	0.14(0.04)***	0.06(0.03)*	0.05(0.05)
W*RMI			-0.04(0.40)	0.56(0.4)
W*DC (log)			41.47(14.11)***	10.68(5.27)**
W*DC^2 (log)			-4.36(1.61)***	-1.19(0.71)*
W*RSI			0.81(1.56)	1.46(1.16)
W*RDI			0.90(3.1)	0.43(2.00)
W*GDP (log)			-7.80(3.38)**	5.78(7.42)
W*GOV			-2.65(0.74)***	-0.30(0.64)
W*RES			-8.20(10.72)	0.59(6.11)
W*OPEN			0.23(0.13)*	0.05(0.07)
Constant	114.96(29.03)***	147.49(63.44)**	200.32(44.74)***	51.68(26.08)*
Adjusted R^2	0.60	0.47		
F-statistic	14.79	7.91		
ρ			0.88(0.40)**	0.91(0.11)***
Log-L				-205.42
Obs.	64	64	64	64

Notes: 1) Significance: 0.01 (***), 0.05 (**), and 0.1 (*).

2) Standard errors of parameter estimates are reported in parentheses.