Spatial attenuation of agglomeration externality:

Evidence from firm-level data in Vietnam

August 5, 2016

Phan Thi Van[#] and Eiichi Tomiura[*]

Abstract

Based on firm-level data in Vietnam, we confirm that the agglomeration spillover from the domestic core attenuates with distance among domestic firms, but find that the attenuation speed significantly accelerates at the former national border among private firms even after four decades since the reunification. Robustness of our main findings is confirmed with historical data of regional population before the North-South division as an instrumental variable. Productivity of all firms, especially foreign-owned firms, is sensitive to the market potential or the presence of other firms.

Keywords: agglomeration; distance; border effect; productivity; firm-level data

JEL Classifications: R12, F23, N95, L22, O18

[#] Graduate School of International Social Sciences, Yokohama National University, Japan, and Faculty of Economics and International Business, Foreign Trade University, Vietnam.
[*] Corresponding author: Faculty of Economics, Hitotsubashi University, Japan. E-mail: e.tomiura@r.hit-u.ac.jp.

1. Introduction

Agglomeration is one of the central topics in spatial economics. Previous literature has established that agglomeration externality attenuates with distance. However, various factors should affect the speed of attenuation. The border effect is critical, as has been extensively analyzed in international economics. Foreign-owned firms are also likely to differ from domestic firms in receiving and generating information spillovers. This paper examines spatial attenuation of agglomeration spillover based on firm-level data in Vietnam, and investigates whether the former national border matters for different ownership types of firms after four decades since the reunification.

As surveyed by Rosenthal and Strange (2004), agglomeration externality has been established as a stylized fact. On the geographic scope, previous studies report that the spillover decays rapidly with distance (e.g. Duranton and Overman 2005). However, the spillover does not attenuate monotonically, affected by borders. On the border effect, German reunification has been actively studied as a natural experiment case. Redding and Sturm (2008) is a prime example of research uncovering persistent border effect. The investigation of such historical division and reunification provides us of a precious opportunity to discuss the impact of border on economic activities. Differences in corporate organization, especially foreign-owned or state-owned compared with privately-owned firms, should also matter in spillovers, but empirical evidence on this aspect has so far been limited.[1]

Vietnam is a suitable case for our research, as the country experienced the division and the reunification in the last century. Vietnam was divided into South and North with the border at the 17 degree latitude after the end of World War II until the reunification of two regions in 1975. The contrast between these two divided countries was clear in economic system: capitalist

[1] For example, Henderson (2003) finds that single-plant firms are more sensitive to externality than multi-plant firms in the U.S. A brief survey of related literature will be given in the next section.

South and socialist North. As the reunification was realized by the victory of North and withdrawal of U.S. from South, state-owned enterprises (SOE) remain powerful in some market segments of this country. As four decades passed since the reunification, however, Vietnam has recently become active in economic reforms since the start of Doi Moi (meaning renovation in Vietnamese) in the mid-1980s. Vietnam became a member of ASEAN and its free trade area in 1995, gained normal access to U.S. market by signing the bilateral trade agreement in 2001, joined World Trade Organization (WTO) in 2007, and singed Trans-Pacific Economic Partnership Agreement (TPP) with developed countries such as U.S., Japan, and Australia in 2015. Foreign multinationals actively established affiliates in Vietnam to seek low wage for export platform purpose in labor-intensive industries. This paper compares firms in North versus South as well as various types of firms (SOE, private domestic firms and foreign-owned firms).

Due to its natural geography of Vietnam as a long country like a Hoteling's linear city, it is natural to discuss spatial attenuation of agglomeration spillover with distance from the core. While Hanoi is the national capital and the political center, Ho Chi Minh has been the dominant economic core in Vietnam. Ho Chi Minh, as a city alone, occupies nearly one-fifth of Vietnam's GDP.

To preview our principal results, the productivity of domestic firms declines with the distance from the domestic economic core Ho Chi Minh, but the attenuation speed significantly accelerates at the former border even after four decades since the reunification among private domestic firms. The robustness of this finding of persistent border effect is confirmed with historical data (regional population before the division) as an instrumental variable. As IV estimates control for regional variations in natural advantages, the persistent border effect is likely to be rooted in second-nature factors. The finding of the persistent border effect especially in private firms rather than SOEs indicates that it is mainly due not to public policy or

regulation but to spontaneously established patterns of private business networks. On the other hand, the productivity of foreign firms, virtually all of which are established only recently, is free from the effect of the former border, but instead sensitive to the market potential and to the concentration/presence of other firms.

The rest of this paper is organized as follows. Section 2 reviews related literature. Section 3 describes Vietnamese firm-level data and historical data used for our analyses. Section 4 explains empirical specifications and reports estimation results. Section 5 adds concluding remarks.

2. Literature review

As this paper links the agglomeration literature with accumulated studies of the border effect, this section briefly reviews both strands of research. This section is not intended to be an exhaustive list of all papers on these topics, but served as a starting point for our empirical analysis.

As surveyed by Rosenthal and Strange (2004), agglomeration has been one of the central research topics in economic geography. Previous studies have shown that the effect declines steeply with distance (e.g. Andersson et al. 2004, Duranton and Overman 2005, Rice et al. 2006, and Soest et al. 2006).[2] However, agglomeration spillover does not monotonically attenuate with geographic distance but is also affected by borders. Henderson (2003) finds that the plant's productivity is influenced by employment activities in the same county but not by those in neighboring counties in the U.S. Many previous studies of knowledge spillovers, such as

[2] Andersson et al. (2004) examine the impact of an exogenous policy change (university relocations) on regional productivity in Sweden, and find that about 75% of spillover occurs within 100 kilometers. Rice et al. (2006) discover that the effect on productivity declines sharply with travel time in U.K. by using historical population weighted by geographic distance as an instrumental variable. Soest et al. (2006) report that the effect on employment dies out quickly with distance in the case of South-Holland. Duranton and Overman (2005) find that localization of industries takes place at small scales below 50 kilometers among U.K. establishments.

Belenzon and Schankerman (2013), Li (2014), and Peri (2005), find that citations to patents decline sharply with distance and are strongly constrained by borders, even by state borders within the U.S.

While previous research often examined the state border in the federated U.S., the national border should be more critical in discussing agglomeration or spillover. The border effect has been repeatedly examined in international trade literature. The case of German reunification has recently attracted attention as a natural experiment in this context. For example, Redding and Sturm (2008) find that the population growth in the cities in Western Germany closer to the East-West German border was substantially lower than other Western German cities after the East-West division. Felbermayr and Gröschl (2014) detect persistent negative border effect on current commodity flows in the case of Union-Confederacy border during the Civil War in the U.S. As a study of deep impact of national border from a different angle, Alesina and Fuchs-Schündeln (2007) find that the German division for 45 years has changed people's preferences, especially fundamental policy stances on state intervention. We examine whether the border between North and South Vietnam, which had been effective for around three decades, has a persistent impact on productivity of firms after four decades since the reunification.

In spite of accumulation of established results on agglomeration, several important issues remain unexplored. Among them, agglomeration spillovers are likely to be affected by corporate organization, such as ownership type. Rosenthal and Strange (2003) compare subsidiaries vs. non-subsidiaries in the effect of agglomeration on births of new local establishments, but detect no clear patterns in the U.S. case.[3] Based on the U.S. census data, Henderson (2003) finds that the productivity of single-plant firms is more sensitive to externality than that of multi-plant

[3] They interpret this inconclusive result as suggesting that "subsidiary status is too rough a measure to capture the influence of a hierarchical corporate structure" (p.387).

firms. In a slightly different context, Crozet et al. (2004) find that foreign firms are strongly attracted to agglomeration of domestic firms in France, possibly as foreign firms compete directly against firms from other foreign countries, benefit from inter-firm mobility of qualified workers from French firms, or expect that French local firms know better about attractiveness of each region.[4]

Although many papers have recently used Vietnamese regional data, none has examined agglomeration (e.g. Brambilla et al. 2012 on household income, McCaig 2011 on poverty, and McCaig and Pavcnik 2014 on labor shift from household business). The border effect is not explicitly discussed either. The use of firm-level data combined with historical province-level population census data differentiates our research from previous studies of Vietnamese regions.

3. Data description

This section describes our data for empirical analysis. We construct a novel dataset by merging two distinct statistics: firm-level data from Enterprise Survey and historical province-level data from Population Census of Vietnam.

Annual Enterprise Survey by General Statistics Office covers all state-owned enterprises, all foreign-owned enterprises, and all private-owned enterprises with not less than ten workers, and randomly sampled 20% of private-owned enterprises with less than ten workers in Vietnam, and collects basic information such as sales, inventory, capital (long-term assets), labor (number of full-time workers), location, and industry codes, as in standard firm statistics in many developed countries. We define production by the sales plus changes in inventory (value at the end minus that at the beginning of the year), while many previous studies depend on sales data in estimating production function. As no data on expenditures on materials are available, we

[4] On the other hand, Kamal (2014) finds the spillover effect is strong among the same ownership type of firms in China.

cannot calculate value-added.[5]

The same survey also identifies firms by ownership types: SOEs (state-owned enterprises), private domestic firms, and foreign-owned firms. In Vietnam, after Foreign Investment Law of 1987, foreign investors are allowed to own 100 percent of shares in all industries except defense-related sectors. As a result of the Doi Moi reform, SOEs are required fiscal autonomy and no longer depend on export subsidies, but rarely privatized. In the statistics, SOE is defined by more than 51 percent of shares owned by state, while foreign firms are those with not less than 49 percent of shares owned by foreigners. All other firms are categorized as private domestic firms.

Geographic locations (address) of firms are also identified in the survey. The entire Vietnam is currently divided into 63 provinces, which we merge to 59 for our research to handle changes in provincial borders during the sample period. The province is the appropriate geographical unit for our research of agglomeration, as few people move across provinces. For example, McCaig (2011) reports that only 2.2% of household heads moved across provincial boundaries in Vietnam. GDP of each province is shown in the map as Figure 1.

This paper uses firm-level data at 2006 to avoid possible noises due to the global financial crisis at 2008. At that year, there are 3,530 SOEs, 111,537 private domestic firms, and 4,031 foreign firms in our sample. We drop firms with production, labor or capital zero or negative from our sample. The summary statistics of firm-level data are shown in Table 1. The definitions of the variables will be given in the next section.

We also exploit historical data for instrumental variables in our regressions. Historical regional data are derived from Vietnam's Population Census, which documents regional population covering all provinces in Vietnam since 1921. As no data on firms or regional GDP

[5] Vietnamese Enterprise Survey either contains no data on workers' hours worked or tangible fixed assets distinguished from long-term assets.

are available for pre-division years, regional population data is the sole practical index to measure local economic activities. We photocopied printed various issues of Statistics Yearbooks at National Library at Hanoi, since historical data are not provided in electronic form. The reasons for choosing regional population at a pre-division year as an instrumental variable will be explained in the next section.

4. Empirical results from firm-level regressions

This section explains our empirical specifications and reports estimation results from firm-level data in Vietnam. We start with the baseline specification to measure the attenuation of agglomeration spillover with distance, and then use historical data as instrumental variables to handle the endogeneity problem. In the last section, we examine how the productivity of a firm is influenced by surrounding regions. In all cases, we pay attentions to the variations across different types of firm ownership.

4.1. Empirical specifications

This subsection explains our baseline specifications for empirical analyses. This paper starts with the following standard Cobb-Douglas production function

$$Q_{jir} = A_{jir} K_j^{\alpha_1} L_j^{\alpha_2}. \qquad (1)$$

The suffix *j* identifies firms, while the industry and the region in which the firm is located are indexed respectively by *i* and *r*. The production, capital, labor and the total factor productivity (TFP) are expressed by *Q*, *K*, *L*, and *A*, respectively. Our main target is estimating whether and how much the firm's TFP varies with the distance from the agglomerated core region. This paper estimates the following log-linear specification:

$$\ln Q_{jir} = \alpha_0 + \alpha_1 \ln K_j + \alpha_2 \ln L_j + \beta_1 \ln DistHCM_r + \beta_2 \ln DistB_r * North_r + \gamma \ln GDP_r + \delta \cdot IND_i + \theta \cdot FirmType_j + u_{jir} \quad (2)$$

TFP is characterized as a function of geographic factors, of which the definitions will be given in detail in what follows. The error term is denoted by *u*. Industry effects are controlled by a vector of the dummies *IND*. There are 87 industries at the two-digit level. We also estimate the same specification separately for private domestic firms, SOEs and foreign firms without firm type dummies *FirmType*. Distinguishing these firm types is critical for current Vietnam as SOE remains powerful in some market segments even after liberalization and market reform while inward FDI is expanding amid globalization. The investigation of differences in the spillover effect across firm ownership types is also informative as a study of agglomeration and economic geography, as no previous research has so far distinguished firm types in estimating spatial attenuation of agglomeration spillover to the best knowledge of the authors.

As the key variable of our interest, the geographic distance of the region from the domestic economic core Ho Chi Minh in kilometers is denoted by *DistHCM*.[6] For firms located within Ho Chi Minh, we measure internal distance using the equation proposed by Redding and Venables (2004): $Dist = 2/3 \cdot \sqrt{Area/\pi}$, where *Area* refers to the area of Ho Chi Minh measured in square kilometers.[7] As the established results from previous literature have confirmed agglomeration externality attenuating with the distance from the center, we expect β_1 be negative.

As we estimate the production function over firms in a cross-section format, however, we should not interpret (2) as indicating the direction of causality. The same equation could be consistent with the self-selection of more productive firms into regions nearer to the congested

[6] We measure the distance from the capital city of each province. Identifying exact address of firms within each province is left for future independent work.

[7] This approximation is based on the average distance between two points in a circular region. The ratio of a circumference of a circle to its diameter is expressed by "π."

Formatted: Justified

core with more intense competition against larger number of rival firms. We do not claim that the estimation of (2) rejects such alternative hypotheses. While Combes et al. (2012) attempt at distinguishing agglomeration from self-selection, this paper focuses on measuring spatial attenuation of agglomeration spillover and examines how it is affected by corporate organization and national border.

To capture the effect of the former national border, we introduce *DistB*, which is defined by the distance from the former national border, the17 degree North latitude, in kilometers. We add this border distance term only for the firms located north of the former border, by interacting with the binary dummy *North*, which takes the value one for firms in the North but zero for those in the South. If the agglomeration externality decays with distance at a higher speed in the former North Vietnam, β_2 should be negative. We also estimate the model with the binary dummy *North* without interacting with *DistB* for a robustness check purpose.[8]

Although Ho Chi Minh is the dominant economic center in this country, the productivity of a firm is likely to be influenced by the size of local market. *GDP* is measured for each province. Positive γ is predicted by the home market effect. By including local GDP in our regression, the agglomeration externality examined in this paper is after controlling for variations in local market sizes.

Although we include it on the right-hand side of the regressions (2), the own region's GDP might not be exogenous. For example, GDP may rise as a result of location of high-productivity firms in the region. GDP in a peripheral region with a limited number of firms may particularly be affected by the productivity of individual firm. To handle this endogeneity problem, we use the following two historical variables as instrumental variables (IV).[9]

[8] The province crossing the former border is excluded when we define *DistB* or *North*.
[9] Even after instruments are assigned to GDP, the potential endogeneity problem cannot be ignored for the firm's choice of production factors *L* and *K*. It is however unfortunately difficult to find an instrument for these variables within our cross-section format. No firm-level data on expenditures on

The first instrumental variable is the regional population at 1943, *POP1943*. The regional population at such an early year can safely be regarded as exogenously given for current productivity of individual firms in the same province but is likely to be correlated with current GDP from the following reasons.

First, population distributions across regions before the division were not affected by North-South national border since Vietnam was a united country at that time. The spatial distribution of economic activities at that time should share some resemblance with current pattern in regions of united Vietnam.

Second, population distribution and distributions of economic activities over regions at such an early year are likely to reflect natural geographic advantages/disadvantages, such as climate, amenity, quality of soils or water, access to the sea or navigable river, and/or abundance of other natural resources. As Vietnamese economy at that time predominantly depended on agriculture or fishery, this assumption appears reasonable. By including population distribution at 1943 in our regression, our estimates can be regarded as those after controlling for natural geography factors.

The other instrument we use for current regional GDP is the bombing intensity during the Vietnam War[10] between 1965 and 1975, *BOMB*. As Miguel and Roland (2011) note, the intensity of bombing varies substantially across regions in Vietnam. The bombing during the war after four decades is exogenous for current firms' productivity but could be correlated with current GDP. If damage of the war, such as remained explosives, for example, still affects economic activities, *BOMB* is negatively related with current GDP of the region.[11] If heavily

materials are available.

[10] In Vietnam, this war is called as "War against U.S." This paper expresses it as Vietnam War to facilitate the understanding for international readers.

[11] The bombing intensity during the war was not systematically linked with the region's population before the war. For example, central regions near the 17th parallel were heavily bombed but the two largest cities, Saigon and Hanoi, were not.

Formatted: Justified

bombed regions recovered from destruction faster than other regions, *BOMB* is instead positively related.[12] The bombing intensity is measured in terms of total U.S. bombs, missiles and rockets per square kilometers.[13]

4.2. Estimation results

4.2.1. Preliminary results

Before reporting the main results, Table 2 presents preliminary OLS results from parsimonious specifications. The dependent variable is the firm's production, as in our main specification explained in the previous section. Industry dummies and firm type dummies are included in all cases. Standard errors are clustered at the province level, as the productivity shocks of firms in the same province are likely to be correlated.

The column (1) is the regression only on the firm's primary factor of production *K* and *L* without any geographic variables. The sum of both coefficients is close to one but slightly less than one, suggesting weakly decreasing returns to scale of production in Vietnam or omitted variable bias possibly due to unavailable data on material inputs or human skills.

The columns (2) and (3) add the distance from the core and the variable for the North-South border. The distance from the former border for firms in the northern locations is used in (2), while the binary dummy for them is used in (3). In both cases, we find that the spatial attenuation of agglomeration externality and the border effect are both strongly significant. The estimated coefficient on the logarithm distance, around –0.1, is in a comparable range with that for knowledge spillovers estimated by Peri (2005).

[12] Davis and Weinstein (2002) conclude that the latter is the case in Japan after the World War II. Miguel and Roland (2011) find no significant impact of bombing during the Vietnam War on the region's poverty and consumption at 2002.

[13] Professor Gérald Roland kindly provided us of his data, which was constructed from U.S. military records and used for Miguel and Roland (2011). This dataset includes wide categories of weapons, except anti-personnel landmines.

The last two columns of Table 2 further add local GDP. While the border effect, both in *DistB* and the binary *North*, remains clearly significant, the distance from the core turns out to be statistically insignificant if we add the province's GDP. As pointed out in the previous section, this result could be contaminated by endogeneity problem. We must wait for IV results before discussing our main topic.

4.2.2. IV results

Table 3 reports the estimation results with IV assigned to local GDP. In the first-stage regression, though omitted from the table, we confirm that we are not affected by the problem of weak instruments, especially by the significantly positive association with historical population.[14] We confirm that local GDP, even after instrumented, remains positive, but, in contrast to Table 2, the first column of this table shows that the distance from the core becomes significant at any conventional significance level. We also confirm that the magnitude of estimated coefficient on *DistHCM* remains basically unaltered from the previous specification without local GDP in Table 2. The distance from the former border for firms in the north remains significantly negative, as in the previous table. The estimates in Table 3 imply that the attenuation of agglomeration spillover significantly speeds up when crossing the former national border at 17th degree latitude.

To explore underlying mechanism behind such persistent border effect, Table 3 further disaggregates firms by ownership types. Instead of adding firm type dummies, the remaining three columns of this table reports regression results separately from each firm type. We

[14] The first-stage OLS confirms that local GDP is highly positively related with population before the North-South division. While GDP of a province declines with the distance from Ho Chi Minh, the relationship with bombing intensity during Vietnam War turns out to be insignificant. We confirm that our main results remain basically untouched whether or not *BOMB* is dropped from the first-stage regression.

emphasize two notable findings from this table.

As the first point to note, the agglomeration externality from the domestic core Ho Chi Minh decays with distance significantly only for domestic firms. Its impact on productivity of foreign firms is only weakly detected at a generous 10% level. Remoteness from the domestic core is serious only for domestic firms.

Table 3 also confirms that the productivity of all types of firms tends to be high in regions with larger local market. For domestic firms, even if we consider this strong home market effect, the distance from the core remains to have a significantly negative impact on productivity.

Several plausible interpretations are in order. First, although no data on exporting of goods are available within our firm-level data set, foreign firms are likely to depend more on foreign sales and thus less susceptible to remoteness from the domestic core market.[15] As a related finding supporting this interpretation, the productivity of foreign firms appears negatively related with the distance from the major international port Hai Phong.[16] Regression results are shown in Appendix Table A1. The impact of the distance from the port on foreign firms is statistically significant when we exclude firms in Hanoi, as foreign firms are likely to be attracted to the political center. We also confirm that the exclusion of Hanoi does not alter our main finding of decaying spillover from the domestic core only for domestic firms. No such relationship is detected for private domestic firms, as most of them are too small to be direct exporters.[17] The significant relation found for SOEs is plausible as they are large in size, capital intensive, and subsidized by the government to export their products. This contrast suggests that

[15] Exporters or importers are not identified in Vietnamese Enterprise Survey. No previous research on Vietnam has used micro-data on trade. McCaig (2011) constructs province-level tariffs from industry-level tariffs weighted by regional employment shares. Brambilla et al. (2012) use the share of fishing income in total household income to analyze U.S. antidumping duty on catfish.

[16] The other large port is located in or near Ho Chi Minh. The distance from the northern border with China is almost automatically inversely related with the distance from Ho Chi Minh.

[17] McCaig and Pavcnik (2014) find that employment shifts away from household business toward registered enterprises especially in provinces nearer to the major ports.

the productivity of exporters is affected by the access to the major international port rather than to the domestic core market.

Second, although we cannot trace knowledge flows between firms, foreign firms can receive knowledge spillover from foreign parent firms, probably located in advanced countries, and employ superior production technologies and/or management know-hows. Hence, either through demand channel or knowledge channel, foreign firms are likely to be less influenced by domestic agglomeration. Although no direct tests for these interpretations are readily available within our limited data, the finding of such a difference between foreign firms and domestic firms in spillover of agglomeration should be worthwhile.

As another important finding from Table 3, the significant border effect is detected only among private firms. The observation of the significant border effect, still powerful even after four decades since the reunification, is driven not by state-owned or foreign-owned firms. As virtually all foreign firms are established only recently, no effect of former border on foreign firms is as expected. Following the policy and central planning by the national government, SOEs are likely to choose production technologies or management practices irrespective of the former border. In other words, such a persistent border effect may be rooted not in exogenous regulations or public policies but probably in private business transactions, customs or culture, often intertwined with natural geography. As displayed in the lower part of Table 3, we also confirm such persistent border effect only for private domestic firms even if the binary dummy *North* is without interacting with *DistB*. We also confirm that our main findings are virtually unaltered even if the firms in Hanoi, the national capital, or in Da Nang, the largest city in the central region of the country, are excluded from the sample.[18]

This effect of distance interacted with the border suggests that agglomeration spillover

[18] Estimation results from these limited samples are available upon request.

decays more seriously after crossing the former border possibly due to barriers in knowledge spillover, or in matching or learning. Institutional difference across the former border may interact with physical distance and dilute the spillover of agglomeration externality to remote locations, or alternatively northern regions near the former border has become assimilated with southern regions through relatively active interactions with firms in the South and receive more benefit from agglomeration. As a related finding of the effect of the distance from the border, Redding and Strum (2008) discover that West German cities closer to the East-West German border experienced a substantial decline in population growth after the division of Germany.

To explore the underlying determinants for such persistent North-South gap, we need to carefully collect more detailed data, for example, on regional characteristics. As Alesina and Fuchs-Schündeln (2007) find in the German case, preference of people and/or fundamental behaviors of workers and managers might have changed under contrasting regimes for three decades of North-South division. Exchanges of goods and services might be relatively inactive across the former border even after long years since reunification, as Felbermayr and Gröschl (2014) report that the North-South border during the Civil War lowers current commodity flows by around 13% in the U.S. The analysis along this line in depth is left for future work, but our preliminary regressions suggest that actual geography, such as heavy costs of transports or travels in mountain areas compared with flat areas along a navigable river, might affect the speed of spatial attenuation of spillovers.[19]

4.3. Cross-regional impacts

Although our main focus of this paper is examining whether the spatial attenuation of

[19] Appendix Table A2 reports the regression results with region dummies, where Vietnam is divided into six broadly defined regions. We find that the negative effect of the distance from the core is particularly large for private firms in Region 2, which is the mountain area in the North.

agglomeration spillover is affected by national border or corporate ownership, this section considers spillovers from regions other than the core or own province. As the previous section shows that the productivity of foreign firms is not affected by domestic agglomeration or former border, we need to explore geographic determinants of productivity for any type of firms, especially for foreign firms. For this purpose, we expand our scope from the sole focus on Ho Chi Minh to the inclusion of all regions in this country. To capture the effects of surrounding regions, we first introduce the market potential *MP* as in Harris (1954) by

$$MP_r \equiv \sum_q \frac{GDP_q}{D_{rq}} \qquad (3)$$

where D_{rq} is geographical distance between capitals of provinces *r* and *q*. This index is a weighted average of regional GDP summed over all provinces with inverse distance as the weight. While market potential has been intensively examined in its impact on wage or firm location, the estimation of its impact on firm-level productivity has been so far limited.[20] As in the previous regressions, GDP of own region is weighted by the inverse of the internal distance. We replace the distance from Ho Chi Minh City and GDP of the own region in (2) by this market potential variable as follows.

$$\ln Q_{jir} = \alpha_0 + \alpha_1 \ln K_j + \alpha_2 \ln L_j + \lambda_1 \ln MP_r + \lambda_2 North_r + \delta \cdot IND_i + v_{jir} \qquad .$$

(4)

As GDP of own region is merged with all other regions, we treat market potential as exogenous for each firm. To capture the border effect, we introduce the binary dummy *North*, which takes

[20] Head and Mayer (2004) find a significant impact of market potential on locations of Japanese affiliates in Europe. Ottaviano and Pinelli (2006) find the strong effect of market potential on productivity at the region level, not firm level in Finland. To estimate the impact of university location on regional productivity in Sweden, Andersson et al. (2004) construct a weighted average of the number of students and researchers of each region with inverse distance as weights, and detect spillovers at the municipality-level.

the value one for firms locating north of the 17 degree latitude and zero for those in the former South Vietnam. The error term is denoted *v*. We estimate (4) with firm type dummies or separately for different types of firms. We expect the coefficient on *MP* be positive.

Although *MP* is defined based on regional GDP, the productivity of a firm may be influenced by the mere presence of other firms. Henderson (2003) shows that high productivity in agglomerated regions is due to the presence of other firms *per se* rather than their large size of production/employment. As firms "could be interpreted as separate source of information spillovers" (Henderson 2003: p.18), the count (rather than the size) of such sources should be related with externality. To capture this effect, following Crozet et al. (2004), we next introduce an alternative measure of spillovers from surrounding regions in terms of the number of firms:

$$NP_r \equiv \sum_q \frac{\#Firms_q}{D_{rq}}. \qquad (5)$$

The number of firms in each province is expressed by *#Firms*.[21] The coefficient on *NP* is expected positive.

Table 4 reports the results with market potential *MP*, while Table 5 presents the results with the presence of other firms *NP*. The binary dummy *North* and industry dummies are included in both cases. As the results in these two tables are similar, we discuss them combined.

As the most important finding from these tables, the spillover from all the regions, either measured in *MP* or *NP*, are significantly positively related with the productivity of all types of firms. Firms located in regions surrounded by larger markets or larger number of other firms tend to be more productive. We also find that the magnitude of its impact, or the elasticity of productivity with respect to *MP* or *NP*, appears the largest for foreign-owned firms. Combined with our previous results on the agglomeration spillovers from the domestic core, the

[21] As in *MP*, we include firms in the same region discounted by the average internal distance.

productivity of foreign firms is likely to be influenced more by demand or the presence of other firms in surrounding locations. From the same tables, we find that the border effect is statistically significant only for domestic firms, not for foreign firms.

This finding of significantly positive impact of market potential on firm-level productivity is consistent with previous results, such as the case of Spain by Holl (2012). In a study of a similar historical natural experiment, Wolf (2007) also finds the significant impact of market potential on industrial reallocations after the reunification of Poland in the early twenty century.

5. Concluding remarks

The use of firm-level data in Vietnam enables us to examine how the spatial decay of agglomeration externality with distance is affected by historical, institutional or organizational factors. The investigation of Vietnam is suitable for this research purpose, since the country experienced the division into North and South and then reunification in the past century, and is now reforming and liberalizing its socialist economy by initiating Doi Moi and by joining WTO and TPP. We have actually found the persistent effect of former national border on spatial attenuation of agglomeration spillovers even after four decades since the reunification among domestic firms, especially private domestic firms. Historical data on population at the province level at 1943 are used as an instrumental variable for local GDP. Such a persistent border effect among private firms appears not to be directed by public policy but instead rooted in perceptions of residents or influenced by actual geography, which are beyond our analysis. The productivity of firms, including foreign-owned firms, is sensitive to the market potential and to the concentration of other firms.

While these findings are informative for understanding the role of history and institution in agglomeration in particular or in economic geography in general, there still remain several

important issues. For example, tracing dynamic entry and exit patterns of firms will help us discuss causality direction. Distributional information of firm productivity, including dispersion or skewness as examined by Okubo and Tomiura (2014) in the Japanese case, will reveal additional richer regularities for discussing agglomeration.

Acknowledgements

Phan appreciates supports by Vietnam government scholarship. Tomiura's research is partly financed by Japan's Grant-in-Aid for Scientific Research C25380284. Professor Gérald Roland kindly provided us with his data on bombing in Vietnam for our research.

References

Alesina, A., and Fuchs-Schündeln, N. (2007) "Good-bye Lenin (or not)? The effect of communism on people's preferences," *American Economic Review* 97(4), 1507-1528.

Andersson, R., Quigley, J., and Wilhelmson, M. (2004) "University decentralization as regional policy: the Swedish experiment," *Journal of Economic Geography* 4, 371-388.

Belenzon, S., and Schankerman, M. (2013) "Spreading the word: Geography, policy, and knowledge spillovers," *Review of Economics and Statistics* 95(3), 884-903.

Brambilla, I., Porto, G., and Tarozzi, A. (2012) "Adjusting to trade policy: Evidence from U.S. antidumping duties on Vietnamese catfish," *Review of Economics and Statistics* 94(1), 304-319.

Combes, P.-Ph., Duranton, G., Gobillon, L., Puga, D., and Roux, S. (2012) "The productivity advantages of large cities: Distinguishing agglomeration from firm selection," *Econometrica* 80 (6), 2543–2594.

Crozet, M., Mayer, T., and Mucchielli, J. (2004) "How do firms agglomerate? A study of FDI in

France," *Regional Science and Urban Economics* 34, 27-54.

Davis, D., and Weinstein, D. (2002) "Bones, bombs, and breaking points: The geography of economic activity," *American Economic Review* 92, 1269-1289.

Duranton, G., and Overman, H. (2005) "Testing for localization using micro-geographic data," *Review of Economic Studies* 72, 1077-1106.

Felbermayr, G., and Gröschl, J. (2014) "Within U.S. trade and the long shadow of the American secession," *Economic Inquiry* 52(1), 382-404.

Harris, C. (1954) "The market as a factor in the localization of industry in the United States," *Annals of American Geographers* 64, 315-348.

Head, K., and Mayer, T. (2004) "Market potential and the location of Japanese investment in the European Union," *Review of Economics and Statistics* 86(4), 959-972.

Henderson, J.V. (2003) "Marshall's scale economies," *Journal of Urban Economics* 53, 1-28.

Holl, A. (2012) "Market potential and firm-level productivity in Spain," *Journal of Economic Geography* 12, 1191-1215.

Kamal, F. (2014) "Does firm ownership affect spillover opportunities? Evidence from Chinese manufacturing," *Journal of Regional Science* 54(1), 137-154.

Li, Y. (2014) "Borders and distance in knowledge spillovers: Dying over time or dying with age? ---Evidence from patent citations," *European Economic Review* 71, 152-172.

McCaig, B. (2011) "Exporting out of poverty: Provincial poverty in Vietnam and U.S. market access," *Journal of International Economics* 85, 102-113.

McCaig, B., and Pavcnik, N. (2014) "Export markets and labor allocation in a low-income country," NBER Working Paper No. 20455.

Miguel, E., and Roland, G. (2011) "The long-run impact of bombing Vietnam," *Journal of Development Economics* 96, 1-15.

Okubo, T., and Tomiura, E. (2014) “Skew productivity distributions and agglomeration: Evidence from plant-level data,” *Regional Studies* 48(9), 1514-1528.

Ottaviano, G., and Pinelli, D. (2006) “Market potential and productivity: Evidence from Finnish regions,” *Regional Science and Urban Economics* 36, 636-657.

Peri, G. (2005) “Determinants of knowledge flows and their effect on innovation,” *Review of Economics and Statistics* 87(2), 308-322.

Redding, S., and Sturm, D. (2008) “The costs of remoteness: Evidence from German division and reunification,” *American Economic Review* 98, 1766-1797.

Redding, S., and Venables, A.J. (2004) “Economic geography and international inequality,” *Journal of International Economics* 62(1), 53–82.

Rice, P., Venables, A., and Patacchini, E. (2006) “Spatial determinants of productivity: Analysis for the regions in Great Britain,” *Regional Science and Urban Economics* 36, 727-752.

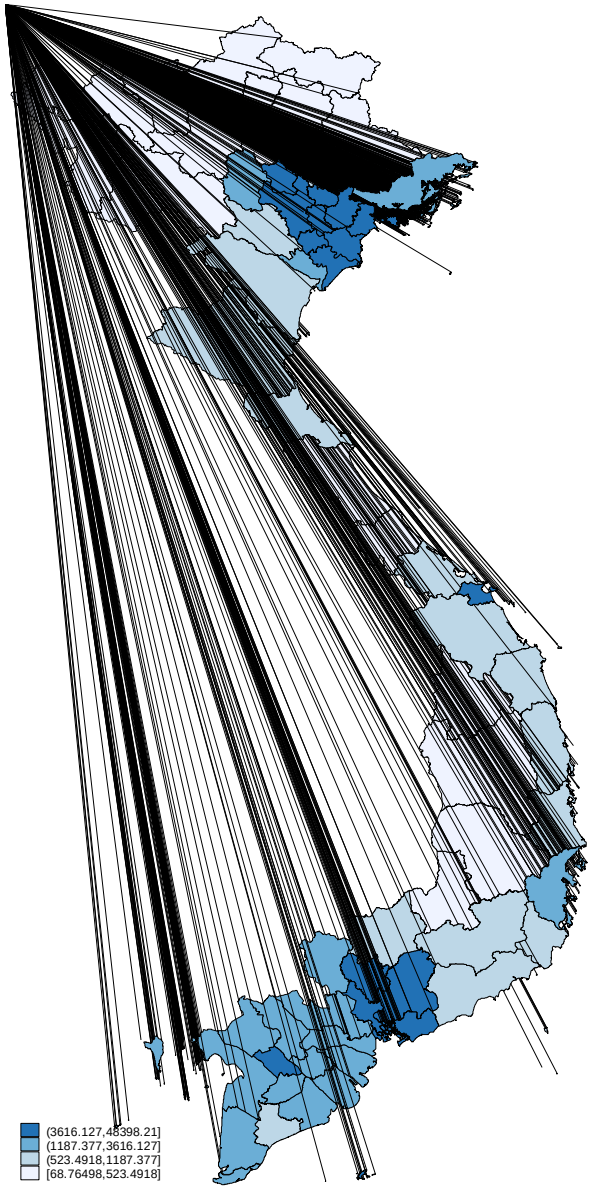
Rosenthal, S., and Strange, W. (2003) “Geography, industrial organization, and agglomeration,” *Review of Economics and Statistics* 85(2), 377-393.

Rosenthal, S., and Strange, W. (2004) “Evidence on the nature and sources of agglomeration economies,” in J.V. Henderson and J.F. Thisse Eds. *Handbook of Regional and Urban Economics* Vol.4, Chapter 49, 2119-2171, Elsevier, Amsterdam.

Soest, D., Gerking, S., and van Oort, F. (2006) “Spatial impacts of agglomeration externalities,” *Journal of Regional Science* 46(5), 881-899.

Wolf, N. (2007) “Endowments vs. market potential: What explains the relocation of industry after the Polish reunification in 1918?” *Explorations in Economic History* 44, 22-42.

Figure 1: GDP of provinces



Note: GDP per square kilometers at 2006 is shown. Truong Sa and Hoang Sa islands are omitted.

Table 1 Summary statistics

		#Obs.	Average	St. Dev	Min	Max
Q	*All*	119,134	23	409	0.001	80,800
	SOE	3,530	269	1,756	0.001	62,000
	Private	111,573	10	72	0.001	15,000
	Foreign	4,031	155	1,426	0.001	80,800
L	*All*	119,134	53	435	1	87,225
	SOE	3,530	483	1869	1	87,225
	Private	111573	29	128	1	10,059
	Foreign	4,031	358	1347	1	55,468
K	*All*	119,134	11	560	0.001	154,000
	SOE	3,530	199	3147	0.002	154,000
	Private	111573	3	78	0.001	21,000
	Foreign	4,031	85	636	0.002	24,100
North		118,507	0.40	0.49	0.00	1.00
DistHCM		59	1,006	752	17	2,111
*DistB*North*		28	630	205	109	1,041
Dist Hai Phong		59	915	727	15	2,012
Local GDP		59	9.1	14.3	0.8	101.0
Pop1943		59	383	333	5	1,294
BOMB		59	29	50	0.01	335
MP		59	2.2	1.5	0.6	8.6
#Firms		59	2,019	4,843	240	32,482
NP		59	520	394	158	2,319

Notes: Production (Q), capital (K), and GDP are in billion VND. Population is in thousand people. Distance is in kilometers.

Table 2 Basic regressions

	(1)	(2)	(3)	(4)	(5)
L	0.670***	0.699***	0.699***	0.700***	0.700***
	(0.0309)	(0.0284)	(0.0285)	(0.0292)	(0.0292)
K	0.270***	0.270***	0.270***	0.271***	0.271***
	(0.00611)	(0.00662)	(0.00661)	(0.00630)	(0.00631)
Dist HCM	--------	−0.0988***	−0.100***	−0.00909	−0.0139
		(0.0234)	(0.0233)	(0.0343)	(0.0337)
DistB North*	--------	−0.0345**	--------	−0.0687***	--------
		(0.0168)		(0.0199)	
North DUM	--------	--------	−0.212**	--------	−0.419***
			(0.105)		(0.123)
Local GDP	--------	--------	--------	0.114***	0.111***
				(0.0224)	(0.0227)
# Firms	119,134	118,507	118,507	118,507	118,507
R^2	0.629	0.653	0.653	0.655	0.655

Notes: The dependent variable is the firm's production. All the variables, except dummies, are in logarithms. Industry dummies, firm type dummies and the constant term are included in all cases, but omitted from the table. Standard errors clustered at the province level are shown in parentheses. Statistical significance is denoted by asterisks: *** at 1%, ** at 5%, and * at 10%.

Table 3 IV results disaggregated by ownership types

	(1) All firms	(2) Private	(3) SOE	(4) Foreign
L	0.724***	0.684***	0.870***	0.850***
	(0.0292)	(0.0308)	(0.0286)	(0.0313)
K	0.281***	0.265***	0.256***	0.307***
	(0.00622)	(0.00753)	(0.0153)	(0.0335)
Dist HCM	−0.0983***	−0.0964***	−0.110***	−0.0815*
	(0.0279)	(0.0290)	(0.0159)	(0.0435)
DistB North*	−0.0360**	−0.0376**	−0.0164	0.0153
	(0.0173)	(0.0180)	(0.0120)	(0.0286)
Local GDP	0.105***	0.0917***	0.127***	0.148**
	(0.0314)	(0.0310)	(0.0414)	(0.0703)
# Firms	118,507	110,979	3,500	4,028
R^2	0.653	0.608	0.762	0.671

	(1) All firms	(2) Private	(3) SOE	(4) Foreign
L	0.723***	0.684***	0.870***	0.851***
	(0.0292)	(0.0308)	(0.0287)	(0.0312)
K	0.281***	0.265***	0.256***	0.307***
	(0.00624)	(0.00754)	(0.0153)	(0.0334)
Dist HCM	−0.0997***	−0.0979***	−0.107***	−0.0838*
	(0.0277)	(0.0288)	(0.0164)	(0.0435)
North DUM	−0.222**	−0.231**	−0.116	0.107
	(0.108)	(0.112)	(0.0778)	(0.182)
Local GDP	0.0989***	0.0848***	0.126***	0.152**
	(0.0307)	(0.0304)	(0.0396)	(0.0695)
# Firms	118,507	110,979	3,500	4,028
R^2	0.652	0.608	0.762	0.671

Notes: The type of firms covered by each regression is shown in the top row of each column. Firm type dummies are added in the first column. The second-stage IV results are shown. Industry dummies and the constant term are included in all cases. All variables, except dummies, are in logarithms. Standard errors clustered at the province level are shown in parentheses.

Table 4 Market potential

	(1) Private	(2) SOE	(3) Foreign
L	0.682***	0.869***	0.850***
	(0.0304)	(0.0270)	(0.0317)
K	0.265***	0.256***	0.307***
	(0.00755)	(0.0146)	(0.0343)
MP	0.181***	0.235***	0.355***
	(0.0383)	(0.0264)	(0.0829)
North	−0.441***	−0.327***	0.118
	(0.0553)	(0.0518)	(0.107)
# Firms	110,979	3,500	4,028
R^2	0.608	0.763	0.672

Notes: Industry dummies and the constant term are included in all cases. All variables including *MP*, except dummies, are in logarithms. Standard errors are clustered at the province level.

Table 5 Presence of other firms

	(1) Private	(2) SOE	(3) Foreign
L	0.681***	0.867***	0.850***
	(0.0305)	(0.0272)	(0.0317)
K	0.266***	0.257***	0.309***
	(0.00767)	(0.0147)	(0.0346)
NP	0.138***	0.199***	0.313***
	(0.0327)	(0.0259)	(0.0699)
North	−0.499***	−0.389***	0.0236
	(0.0515)	(0.0482)	(0.0750)
#Firms	110,979	3,500	4,028
R^2	0.607	0.763	0.673

Notes: See notes to Table 4.

Appendix

Table A1 Distance from the major international port

	(1) Private	(2) SOE	(3) Foreign	(4) Private	(5) SOE	(6) Foreign
L	0.680***	0.870***	0.855***	0.653***	0.865***	0.842***
	(0.0306)	(0.0289)	(0.0314)	(0.0175)	(0.0384)	(0.0319)
K	0.266***	0.259***	0.301***	0.273***	0.257***	0.296***
	(0.00788)	(0.0152)	(0.0330)	(0.00854)	(0.0213)	(0.0365)
Dist Hai Phong	−0.0178	−0.0819***	−0.0561	−0.0210	−0.0826**	−0.127***
	(0.0830)	(0.0268)	(0.0559)	(0.0860)	(0.0349)	(0.0467)
South	0.592***	0.620***	0.310	0.613***	0.702***	0.681***
	(0.213)	(0.110)	(0.222)	(0.214)	(0.126)	(0.209)
Local GDP	0.104***	0.101**	0.122**	0.0965***	0.0853*	0.101*
	(0.0228)	(0.0378)	(0.0516)	(0.0249)	(0.0448)	(0.0531)
Regional coverage	All	All	All	Excluding Hanoi	Excluding Hanoi	Excluding Hanoi
# Firms	110,979	3,500	4,028	91,245	2,754	3,572
R^2	0.606	0.760	0.670	0.606	0.767	0.668

Notes: IV results at the second stage are shown. *South* is the binary dummy for firms in the south. Industry dummies and the constant term are included. Standard errors are clustered at the province level.

Table A2 Variations across regional blocks

	(1) Private	(2) SOE	(3) Foreign
L	0.694***	0.868***	0.854***
	(0.0311)	(0.0274)	(0.0300)
K	0.262***	0.254***	0.299***
	(0.00674)	(0.0143)	(0.0323)
DistHCM	−0.0110	−0.110***	−0.00661
	(0.0322)	(0.0280)	(0.0419)
*DistHCM*R1*	−0.0718***	0.00160	−0.0147
	(0.0201)	(0.0148)	(0.0222)
*DistHCM*R2*	−0.126***	−0.0442**	−0.120***
	(0.0211)	(0.0175)	(0.0283)
*DistHCM*R3*	−0.0717***	−0.0227*	−0.0504**
	(0.0183)	(0.0128)	(0.0242)
*DistHCM*R4*	−0.0503**	−0.0119	−0.200***
	(0.0208)	(0.0242)	(0.0471)
*DistHCM*R6*	0.00698	0.0408**	−0.0179
	(0.0185)	(0.0169)	(0.0237)
Local GDP	0.0783***	0.0834**	0.183***
	(0.0285)	(0.0355)	(0.0411)
# Firms	111,573	3,530	4,031
R^2	0.613	0.764	0.676

Notes: Regional dummies are denoted by *R1, R2,…*, and *R6*. The second-stage IV results are shown. Industry dummies are included in all cases. See also notes to Table 3. Vietnam is divided into the following six regions:

R1: Red river delta: Ha Noi, Ha Tay, Vinh Phuc, Bac Ninh, Quang Ninh, Hai Duong, Hai Phong, Hung Yen, Thai Binh, Ha Nam, Nam Dinh, Ninh Binh.

R2: Northern midlands and mountainous: Ha Giang, Cao Bang, Bac Kan, Tuyen Quang, Lao Cai, Yen Bai, Thai Nguyen, Lang Son, Bac Giang, Phu Tho, Dien Bien, Lai Chau, Son La, Hoa Binh.

R3: North Central and South Central Coast: Thanh Hoa, Nghe An, Ha Tinh, Quang Binh, Quang Tri, Thua Thien - Hue, Da Nang, Quang Nam, Quang Ngai, Binh Dinh, Phu Yen, Khanh Hoa, Ninh Thuan, Binh Thuan.

R4: Highlands: Kon Tum, Gia Lai, Dak Lak, Dak Nong, Lam Dong.

R5: South East: Binh Phuoc,Tay Ninh, Binh Duong, Dong Nai, Ba Ria Vung Tau, Ho Chi Minh.

R6: Mekong Delta: Long An,Tien Giang, Ben Tre,Tra Vinh,Vinh Long,Dong Thap,An Giang, Kien Giang, Can Tho, Hau Giang, Soc Trang, Bac Lieu, Ca Mau.