

Does Tokyo Matter?

Increasing Returns and Regional Productivity

by

Donald R. Davis
Harvard University
Federal Reserve Bank of New York
and NBER

and

David E. Weinstein
University of Michigan
and NBER

Preliminary: Draft 0.1

June 1999

Abstract: One account of spatial concentration focuses on productivity advantages arising from market size. We investigate this for forty regions of Japan. Our results identify important effects of a region's own size, as well as backward linkages between producers and suppliers of inputs. Productivity links to a more general form of "market potential" or Marshall-Arrow-Romer externalities do not appear to be robust in our data. Landlocked status does not matter for productivity of regions in Japan. The effects we identify are economically quite important, accounting for a substantial portion of cross-regional productivity differences. A simple counterfactual shows that if economic activity were spread evenly over the forty regions of Japan, aggregate output would fall by nearly twenty percent.

We have benefitted from excellent research assistance by Paris Cleanthous, William Powers, and Pao-Li Chang. Gordon Hanson graciously provided us with a program for calculating great arc distances. The views expressed in this paper are those of the authors and do not express any official view of the Federal Reserve system.

Correspondence to David E. Weinstein, University of Michigan Business School, Ann Arbor, MI 48109

Phone: Davis (212) 720-8401, Weinstein (734) 936-2866, e-mail: ddavis@fas.harvard.edu, weinstei@umich.edu

Does Tokyo Matter? Increasing Returns and Regional Productivity

I. Disparity in Regional Economic Density

The spatial distribution of economic activity is extraordinarily uneven. Hyper-dense cities exist, as do large regions of sparse activity. For example, in the Japanese regional data with which we will work, GDP per square kilometer varies across regions by a factor of twenty-two. This vast disparity in regional economic concentration is the central stylized fact motivating a recent theoretical revival in the areas of regional and urban economics. Many of the most important contributions to this revival are summarized in *The Spatial Economy*, by Fujita, Krugman, and Venables (1999).

The spectacular variation in regional economic density is indisputable. The key questions, though, are two: (1) Why does the regional variation arise? and (2) How much does it matter? The theory in this area is fertile, generating a large number of hypotheses. We would like to distinguish two broad accounts, which may be complementary. A first account focuses on the advantage to consumers of locating in a large market. Such an account typically is developed in a Dixit-Stiglitz framework, with an interaction between a preference for variety and product-level scale economies. An important feature of the most commonly-used Dixit-Stiglitz specification is that the equilibrium degree of product-level economies of scale is invariant to region size. That is, larger regions yield benefits to consumers in this case not because of enhanced productivity in an input-output sense, but rather indirectly through a variety-influenced decrease in the price index.

The second account of disparity in regional economic density focuses on productivity advantages that arise from economic size. There are many channels through which size may matter. These include externalities, whether purely local or mediated by distance, and market access, affecting both forward and backward linkages. The precise channels and the measures we will use to implement them will be reviewed in the following sections. What we would like to emphasize at this point is that if these stories matter substantively for welfare, then we should be able to verify that having a large market in the relevant sense raises average productivity in a region.

This paper aims to identify the channels by which market size may influence productivity and measure the strength of this influence. We examine this in a sample of 40 Japanese regions, utilizing the same data as Davis, Weinstein, Bradford, and Shimpo (1997), Davis and Weinstein (1998), and Bernstein and Weinstein (1998). The premise for our study is that cross-regional variation in average productivity will have observable implications for the relation between the national technology, regional output, and regional factor supplies. Since theory provides many accounts, we look to the data to identify which seem most important.

II. Theory

The objective of this section is to develop a regional economic geography model with free factor mobility and trade between regions in which it is possible to explore the relation between region size and productivity. The analytic work draws strongly on previous work in economic geography, as summarized in FKV (1999). However, it also introduces a few tricks of its own

that simplify the problem so that we may focus on the aspects that are crucial here. It is simplest to start with a stripped-down model and add complexity as needed.

A. Market Size and Productivity

We begin with a specification of technology due to Ethier (1979). In each region, there is a single homogeneous final output, given by Y . Production of the final output Y draws on a continuum of intermediate inputs of measure n . If $q(i)$ is the usage of input i , then the production function is:

$$Y = \left[\int_0^n q(i)^\rho di \right]^{1/\rho} \quad \text{where } 0 < \rho < 1$$

Since the final output producer takes n as given, this gives rise to a cost-minimizing demand for the typical intermediate as:

$$q(i) = \frac{p(i)^{-\sigma}}{\int_0^n p(j)^{1-\sigma} dj} Y$$

The production and market structure for the intermediates are the standard Dixit-Stiglitz monopolistic competition. Hence each producer is a monopolist in the production of its own intermediate. Since that producer takes both aggregate expenditure and the prices of other intermediates as given, it sees itself as facing a demand curve with constant elasticity $\sigma > 1$, where $\sigma = 1/(1-\rho)$. Production of each intermediate takes place under increasing returns to scale. The unit labor requirements to produce $q(i)$ of intermediate input i are given as:

$$l(i) = F + cq(i)$$

With constant elasticity demand, prices of all intermediates will be a simple markup on marginal labor cost:

$$p(i) = wc / \rho$$

Free entry insures that profits are driven to zero. With a fixed markup over marginal cost, this implies that every intermediate producer operates at precisely the scale that allows recovery of fixed costs. As is well known, this gives rise to equilibrium scale and labor usage per variety of:

$$q^* = F(\sigma - 1)/c \quad \text{and} \quad l^* = F\sigma$$

We assume in this first model that there are many regions across which labor may move, but that neither final nor intermediate products may be traded. Let L_r be the labor force in region r . Then the number of intermediates produced in r in equilibrium will be $n_r = L_r / l^*$. With equilibrium scale fixed at q^* , final output is given as:

$$Y_r = n_r^{1/\rho} q^* = L_r^{1/\rho} q^* l^{*-1/\rho}$$

The important thing to note here is that even though the final goods producer perceives itself as producing under constant returns to scale, the elasticity of final output with respect to the regional labor force exceeds unity. The reason, of course, is that there are productivity gains associated with the variety that increases with market size. Noting that q^* and l^* are constant and common across regions, we can convert this to a region-specific labor productivity measure:

$$y_r \equiv Y_r / L_r = L_r^{(1-\rho)/\rho} q^* l^{*-1/\rho}$$

Hence labor productivity rises with region size. If this were all there is to the story, the freely-mobile labor would all pour in to a single region, where it would achieve maximal utility. There are a variety of ways to counteract the centripetal forces we have put in place. One is to specify that "farmers" constitute an immobile source of demand in other regions. Here this would tend to add clutter. A second approach is to specify a congestion technology in some non-traded good such as housing. We pursue a third alternative, which also relies on congestion, but which assumes that the congestion affects utility directly. In particular, we specify that utility for the representative individual in region r is given as:

$$U_r = g(L_r)y_r$$

We further assume that $g'(L_r) < 0$, reflecting utility costs of congestion. It is straightforward to see that such a specification may allow for the co-existence of regions of highly varying size so long as the centripetal and centrifugal forces balance. The intuition comes through quite clearly if we specify the (improbable) case in which: the congestion takes the form:

$$g(L_r) = L_r^{-\frac{1-\rho}{\rho}}$$

In this case:

$$U_r = g(L_r)y_r = q * l^{*-1/\rho}$$

Here the congestion effects exactly offset the productivity advantages of a larger regional market at any region size. As a consequence, every region size is consistent with equilibrium. It is clear from this example that it would be straightforward to specify a congestion function that may give rise to multiple stable region sizes.

B. Market Potential and Backward Linkages

So far our model has not allowed for trade across regions. However, Krugman and Venables (1995), among other papers, has emphasized the important role in productivity that may be played by market access, particularly so-called backward linkages to suppliers of inputs produced under economies of scale.

There is an extended discussion of the role of backward linkages in FKV. The discussion is developed in a setting in which trade costs are explicitly incorporated. For our purposes, though, it will prove useful to take a shortcut. The shortcut is simply to reinterpret the model we have already developed as a model of interregional trade and to focus on the differential productivity advantage conferred by excellent market access.

Consider a world in which trade costs between regions take on just two levels -- zero and infinity. Let L_r be the total labor force of a set of regions $r \in R$, and $L_{r'}$ similarly be the total labor force for regions $r' \in R'$. We assume that trade costs are zero between any pair of regions in R , and similarly for any pair of regions in R' , while trade costs between any pair across R and R' are infinite. For definiteness, let the labor force L_r exceed that of $L_{r'}$ by a factor $\phi > 1$. Then, by our previous reasoning, labor productivity in any region of R exceeds that of any region in R' by a factor $\phi^{(1-p)/p} > 1$. Note that we have stated this without reference to the size of individual regions within R or R' . For example, this holds in comparing a small region in R with a large region in R' . What matters is not own-market size, but rather market access.

We have developed this for the case of a single industry. One interpretation of this would say that all industries can draw on the productivity advantage conferred by access to this broad range of inputs. This would lead us toward a "market potential" approach, as in Harris (1954). Nonetheless, it would be reasonably straightforward to extend our example to a case where industries draw on inputs from some industries to a greater extent than others. From a theoretical standpoint, this offers little additional insight. However, it suggests looking more closely at the precise nature of the linkages between producers and their suppliers. This leads us toward a "backward linkage" approach to regional integration and productivity.

III. Towards Empirical Implementation

A. Study Design

Our study investigates determinants of regional productivity. The dependent variable "productivity" will be described below. We relate productivity to a variety of traditional variables as well as introducing new variables that stress the role of forward and backward linkages.

A first set of variables consists of various measures of market size. The simplest is "own-size," which in this draft will be represented alternatively as the regional labor force or the regional GDP. A variety of rationalizations of why this may affect productivity may be offered. One is that local economic activity gives rise to a pure Marshallian externality. A second is that the variety link to productivity developed in the theory section is very general, so that productivity depends on the level of local activity, but not directly on its composition. An alternative to "own-size" is what Harris (1954) termed "market potential." The latter is a more general framework, which allows productivity to be affected by a weighted average of GDPs of the region itself as

well as its neighbors, where the weights are inverse to bilateral distance. In this sense, the "own-size" variable is one of market potential where all of the weight is placed on local regional output.

Two new variables may be considered, which likewise emphasize issues of market access, but which focus more directly on the linkages between suppliers, users, and final consumers. The variable "backward linkage" emphasizes that productivity may be high when there is excellent access to sources of precisely the inputs required for that particular region's output. This emphasizes the structural input-output links between producers and their suppliers, as discussed in the theory section. One may also consider the structural relation embodied in "forward linkages." One interpretation of the forward linkages suggests that this may matter greatly for location decisions as producers seek to be near their source of demand. However, under this interpretation, there need not be any direct link to productivity. An alternative interpretation, however, might suggest that producers have a great deal to learn from consumers of their product, so that strong forward linkages may also be a source of productivity advantage.

We will also consider two variables which have figured prominently in previous studies. The first is a measure of regional specialization. Glaeser, et al. (1992) examine the role of Marshall-Allen-Romer (MAR) versus Jacobs externalities in city growth. In their schema, the MAR view posits that learning should be greater where there is a concentrated output structure, whereas Jacobs emphasized potential benefits of a diverse production structure. Glaeser, et al. find evidence they interpret as favorable to the MAR view. Our study differs theirs in that it considers the level of productivity rather than city growth. However, if productivity gains are believed to be the source of the differential city growth, then we should be able to find some evidence of this in the resulting productivity levels. We will also examine a suggestion from

Gallup, Sachs, and Mellinger (1998) that landlocked status matters for growth. While their study emphasizes the link to growth in a cross-national study, we will examine whether this extends to productivity for a cross-regional sample. This may provide insight to whether it is the remoteness typical of landlocked regions that matters or the fact that access to the sea must cross national political boundaries.

B. Data Construction

In this section we provide an overview of the data used in the paper. Details on the construction of variables are in the appendix to Davis and Weinstein (1998). Our data set contains output, investment, consumption, government expenditure, endowment, and absorption data for the 47 prefectures/cities of Japan. We form two aggregates: *Kanto*, out of the city of Tokyo and the prefectures of Ibaraki, Kanagawa, Chiba, and Saitama; and *Kinki*, out of the prefectures/cities of Hyogo, Kyoto, Nara, and Osaka. This reduces our sample to 40 observations, but reflects the high level of integration of the prefectures surrounding Tokyo and Osaka. Our distance data is derived from the Kei/Ido Ichiran Database, which provides longitude and latitude data for Japanese cities, allowing calculation of the great arc distance between points.

Define X_r as the $N \times 1$ gross output vector for region r , and $[1]$ as an $N \times 1$ vector of ones. Let AX_r , C_r , I_r , and G_r be prefectural intermediate input demand, consumption, investment, and government expenditure vectors. Construction of these variables is described in more detail in Davis and Weinstein (1999). Define X_r^{TRAD} to be equal to X_r for all manufacturing, agricultural, and mining sectors and zero otherwise. Finally we set $DIST_{rr'}$ equal to the distance

between the prefectural capital cities when $r \neq r'$ and equal to the square root of the area of the prefecture divided by π otherwise.

We now turn to the construction of our key variables. We begin with the measure of productivity, which will be the dependent variable in our study. Previous papers, such as Sveikauskas (1975), Henderson (1986), and others, have looked at productivity differences by estimating regional production functions for particular industries. The standard approach involves either calculating TFP using index numbers or estimating a regional production function. One of the problems with this approach is that it is impossible to identify forward and backward linkages using a production function approach because one needs to have information about the regional availability of inputs and absorption of output. Such information is available if one turns to input-output data.

In this paper we will measure factor productivity using the matrix of direct factor input requirements. Our measure of regional productivity of factor f is π_{rf} where we arbitrarily set the productivity of each factor for Japan as a whole equal to unity. For each region and factor, the following condition must hold:

$$B_f X_r = \pi_{rf} V_{rf}$$

Note here that B_f is the Japanese average input requirement, so unlike the other variables is not specific to region r . Hence, we define productivity in region r of factor f as:

$$\pi_{rf} = \frac{B_f X_r}{V_{rf}}$$

We now turn to specification of our independent variables. Our "Own-Size" variable will measure aggregate regional size, and will be implemented alternatively as the regional labor force

or regional GDP. An alternative measure of a region's size takes account of its proximity to other regions. Following Harris (1954), we define "Market Potential" for region r as:

$$MP_r = k_M \sum_{r'} \frac{GDP_{r'}}{DIST_{rr'}}$$

where

$$k_M = \left[\sum_{r, r'} \frac{GDP_{r'}}{DIST_{rr'}} \right]^{-1}$$

In this definition, as well as in all of our subsequent definitions of variables involving distance, we assume that a one percent increase in distance causes the impact of output of demand to fall by one percent. This choice is based on the typical coefficient obtained in gravity model using both regional and international data.

When we say that a region has strong backward linkages, we mean that it has excellent access within the region and in neighboring regions to the investment goods and intermediate inputs used intensively by that region's producers. An empirical implementation of this concept defines "Backward Linkage" as follows:

$$BACK_r = k_B \left([1]^T (AX_r + I_r) \right)^{-1} \sum_{r'} \frac{[AX_r + I_r]^T X_{r'}^{TRAD}}{DIST_{rr'}}$$

where

$$k_B = \left[\sum_r \left(([1]^T AX_r)^{-1} \sum_{r'} \frac{[AX_r]^T X_{r'}^{TRAD}}{DIST_{rr'}} \right) \right]^{-1}$$

This variable is an input-weighted average of production across all of Japan. Hence backward linkages are strong when the producers of our inputs are large and proximate. This definition only allows backward linkages to occur through tradable goods sectors. This is based on work by Polenske [AER 1963?] verifying the gravity model for Japanese regions and Bernstein and Weinstein (1998) that the size of traditional non-tradable sectors seems to be linearly related to endowments in Japan.

In addition to these core variables, we define a number of other variables that have been used in previous studies. Glaeser et al. (1992) test for the existence of MAR or Jacobs externalities using an index of specialization based on the concentration of employment in particular industries. We will also allow for these factors by following their definition but will use output instead of employment as our measure of concentration. Our measure of specialization is

$$\text{SPECIALIZATION}_r = \frac{1}{40s_r^2} [X_r]^T [\text{Diag}(X_{Japan})]^{-1} [\text{Diag}(X_{Japan})]^{-1} [X_r]$$

If each region were a one fortieth scaled down version of Japan as a whole, then this index would always equal unity. However, as regions concentrate in particular sectors, then this index will be larger.

Finally, we also define a variable that can capture forward linkages. We set forward linkage to be

$$\text{FOR}_r = k_F \left([\mathbf{1}]^T X_r^{\text{TRAD}} \right)^{-1} \sum_{r'} \frac{[X_r^{\text{TRAD}}]^T (AX_{r'} + C_{r'} + I_{r'} + G_{r'})}{\text{DIST}_{rr'}}$$

where k_F is set so that this variable equals one when sum across all prefectures. Our forward linkage variable gives us an output-weighted average of demand across regions. Paralleling our

backward linkage variable, our forward linkage variable is large when the demanders of our tradable goods are large and close.

Table 1 presents sample statistics for all of our variables. There are a number of points that are worth noticing. First, the average deviation in productivity across prefectures is not necessarily zero because small prefectures may have higher or lower factor productivity than large prefectures. This explains why the average deviation is negative for both labor factors and positive for capital. Second, there appears to be more variation in labor productivity than in capital productivity. This may reflect the relatively high degree of capital mobility across Japan. Third, all of our geographic market variables -- market potential, backward linkage, and forward linkage -- are highly correlated. This makes it difficult, though not impossible, to separate the effects of these variables.

C. Estimation Issues

When we move to a multi-factor, multi-good, multi-distance world, analytic solutions become infeasible. Therefore we need to abstract to some degree from the theory in the implementation, while hoping to capture its salient insights. Using our definition of productivity, we can estimate the effects of backward linkages, forward linkages and market potential on productivity through variants of the following equation:

$$(1) \quad \pi_{rf} = \alpha_f + \beta_{1f} MP_r + \beta_{2f} \ln(\text{GDP})_r + \beta_{3f} \text{BACK}_r + \varepsilon_{rf}$$

This gives us one equation for each factor or three equations in total.

There are other estimation issues we need to address. First, the ε_{rf} 's are likely to be correlated across factors since neutral technical differences will affect all factors equally. This

suggests that we should not assume that $\text{corr}(\varepsilon_{rf}, \varepsilon_{if'})$ equals zero. We solve this by adopting an iterative seemingly unrelated regressions technique. Second, it is unlikely that the impact of market size variables should differ across factors. Rather it seems more reasonable that the economic geography variables should have common effects for all factors. We can impose this on the data by forcing $\beta_{rf} = \beta_{if'}$ for each factor, and estimating the set of three equations simultaneously. Finally, we are likely to measure average productivity more accurately in larger regions than in smaller regions because mismeasurement of output and endowments is likely to fall. We therefore weight all observations by the square root of the regional labor force before estimation.

IV. Data Preview and Results

A. Data Preview

Before proceeding to a formal data analysis, it will prove useful to preview certain features of the data. A first issue worth addressing is the level of aggregation used in the analysis. A check on this comes in the form of Zipf's law, an extremely robust feature of national data sets. Zipf's law holds that the log of region size will fall one-for-one with the log of the rank of a region's size. Figure 1 examines this for our Japanese regions. As the plot reveals, this relationship holds almost exactly for Japanese prefectures under our aggregation scheme. The slope coefficient is -0.951 . This reflects the fact that the size distribution of regions is quite skewed. The largest region, Kanto, is about 77 times larger than the smallest region, Tottori. The three largest regions – containing the cities of Tokyo, Yokohama, Osaka, and Nagoya – produce nearly half of Japanese GDP.

Japanese region-size seems also to be positively correlated with our measure of productivity. In Figure 2 we plot the average factor productivity in a region against region size. These variables are clearly positively related. Doubling region size is associated with productivity rising by about 5 percent. This positive relationship between region size and productivity has been confirmed econometrically in a large number of previous studies (e.g. Sveikauskas (1975) and others).

Average productivity of Japanese regions ranges from 27 percent below the national average in Okinawa to as much as 15 percent above the national average in Aichi. These extreme points are quite suggestive of the role that geography may play in regional productivity. Okinawa is not the smallest Japanese prefecture, indeed it is not even in the smallest decile, but it is by far the most remote prefecture, situated about 500 hundred miles Southwest the Japanese archipelago. Shimane prefecture, a more centrally located prefecture with a similar population, has a productivity gap that is only half that of Okinawa's. This is suggestive of the possibility that Okinawa may be at a disadvantage because of it's distance from the mainland.

Hokkaido and Fukuoka are also significant outliers. Despite being the fourth and fifth largest prefectures in Japan in terms of labor force, their productivity is significantly below average. Both of these prefectures are located off the main Japanese island at the Northern and Western extremes and are therefore quite remote from other sources of supply. At the other extreme is Aichi, which has the highest productivity in all of Japan. Aichi contains the moderately-sized city of Nagoya and is only one fifth the size of Kanto and less than one half the size of Kinki. However, situated almost equidistantly between the two largest Japanese regions

on the major Japanese rail lines and highways, producers in Aichi have easy access to goods produced in either of these large regions.

This anecdotal evidence suggests that that we also explore how market access affects productivity. In Figure 3 we plot productivity against our backward linkage variable. Allowing remoteness to matter, we now find that the most productive prefecture, Aichi, has the strongest backward linkages, and the least productive prefecture, Okinawa, has the weakest. The only really troubling point in this plot is Gifu, the second point from the right. Gifu appears to have substantial market access but low productivity. One reason for this is that Gifu's population is 25 percent below that of the average region. A second reason is that Gifu's excellent market access is an artifact of the way we construct the backward linkage variable. For almost all prefectures, the capital city lies in the center of the prefecture. Gifu, however, lies just above Aichi, and since the city of Gifu is only about 20 km from Nagoya, in our data Gifu is closer to Aichi than it is to itself! That is, our measure overstates the strength of Gifu's market access. We could have aggregated Gifu with Aichi or recalculated the backward linkage variable to improve the fit, but we preferred not to change our data construction method in order to eliminate outliers.

B. Results

Table 2 presents the results from estimating equation 1. As is suggested by Figures 2 and 3, there is a strong positive relationship between region size and productivity as well as between region market access and productivity. This relationship is present regardless of whether the variables are considered separately or together. Our estimates indicate that a doubling of region size causes productivity to rise by about 3.5 percent. This we attribute to a pure Marshallian externality.

Of more interest is the role played by market access. Okinawa has a population that is 10 percent larger than Yamanashi (located adjacent to Tokyo). Okinawa has a productivity level that is 27 percent below the national average while Yamanashi is almost exactly at the national average. Our estimates indicate that 10 percentage points of the gap between the two prefectures is due to the greater distance between Okinawa and the mainland. Similarly, Shizuoka prefecture, located just west of Kanto has a slightly smaller population than Hokkaido, but significantly better market access to Kanto, Kinki, and Aichi. Our estimates suggest over half of the 19 percent productivity gap between Hokkaido and Shizuoka is due to the latter's advantage in market access.¹

In Table 3, we conduct a number of robustness tests. Glaeser et al. (1992) include a variable for regional specialization in their growth regressions and find that the coefficient is positive indicating that regions that are more specialized in particular sectors have higher growth rates. They interpret this as evidence in favor of Jacobs' externalities. In the cross-section, one should also expect that specialization should have an impact on productivity. In the first column of Table 3 we include a variable that increases with regional specialization. When included with GDP, we obtain the predicted positive coefficient, indicating that we also find in our data the same positive relationship between productivity and specialization. However, when we control

¹ Clearly size and geography play important roles in understanding the regional distribution of national welfare. This has implications for European integration too. The European Union has a population that is just over twice that of Japan and most European nations have populations that are smaller than Kanto (approx 35 mln) or Kinki (approx 17 mln). Economic geography suggests that countries located near the major economies are likely to be the major winners from integration. Our analysis suggests that Denmark stands considerably more to gain than Portugal.

for backward linkages, we find that the specialization variable ceases to be significant. This suggests that specialization is not that important if one controls for market access.

A number of authors, e.g. Gallup, Sachs, and Mellinger (1998), have suggested that access to the sea is important in understanding regional growth. Although Japan is an island nation, six Japanese prefectures are landlocked. To see whether that mattered, we also included a dummy variable that was one for each of these prefectures. Our results suggest that being landlocked does not have much of an effect on productivity in the affected regions of Japan.

Since regional output is a factor in determining both productivity and backward linkages, we must be concerned about possible simultaneity problems. We intend to deal with this more thoroughly in the next draft, but here are some simple things that we did to see if the backward linkage effect arises from higher than expected domestic output. First, we replaced log of the labor force with log of regional GDP to see if backward linkages were just capturing high production in regions with high productivity. As one can see in Table 4, this had almost no impact on any of the coefficient estimates.

Theory is ambiguous about the role that forward linkages may play in productivity. Clearly in a world with trade costs, it is advantageous for producers to locate near important sources of demand in order to minimize trade costs. However this need not confer on them any productivity advantage in the link between inputs and outputs. Yet this could arise if excellent access to consumers of your product yields information that allows productivity gains. This suggests adding forward linkages to the horse race over how market size matters. We see in Table 1 that forward and backward linkages are highly correlated with each other (as well as market potential), so it will be interesting which the data identifies as key in influencing productivity.

As we noted, this variable is highly correlated with backward linkages ($\rho = 0.95$), so multicollinearity is likely to be a major problem. As we see in Table 5, the addition of the forward linkage variable does increase the standard errors on the coefficient on backward linkages, but the effect that we have identified seems clearly to flow through backward linkages and not through forward linkages.

C. Counterfactuals

While it seems clear that economic geography matters for the distribution of income across regions, the implications for aggregate income are less clear. The importance of backward linkages can be understood by considering a counterfactual experiment. Suppose that one divided the country into 40 equal sized prefectures each with endowments equal to 0.025 of the total and then banned trade between these prefectures.² Our estimate based on the Marshallian externality suggests that there would be a 13 percent drop in Japanese aggregate output arising from concentrating endowments in large regions that can realize scale economies. Shutting off the channel for backward linkages, however, would have a somewhat smaller effect. If demand remained constant, our estimates indicate that productivity would fall by an additional 3 percent due to the fact that large regions tend to be clustered together and benefit from each other's supply of intermediates. If we assume that this sixteen percent drop in productivity would

² This counterfactual is not entirely without historical basis. The modern prefectural system was created when, following the Boshin Civil War (1868-1869), two powerful and victorious feudal lords, one in Kagoshima prefecture (then called Satsuma Domain) and one in Yamanashi prefecture (then called Choshu Domain) voluntarily surrendered their domains to the emperor. Since Satsuma and Choshu had just defeated the Shogun and largely controlled the emperor, this was not a complete abdication of power. The 259 other smaller domains were then aggregated into the modern prefectures. Our counterfactual basically asks what would have happened if these daimyo had not given up their lands and Japan had returned to its pre-shogunal state of hostile domains. (This is not as crazy as it sounds since many in Satsuma soon came to regret their decision and mounted a force of 40,000 against the new government in the doomed Satsuma Rebellion of 1877)

produce a 16 percent drop in aggregate demand, we would observe another 2.5 percent drop in productivity due to the fact that each prefecture's size would shrink by an additional 16 percent. If we iterate in this manner ad infinitum, we find that in equilibrium, output and demand would fall by just over 19 percent.³

V. Conclusion

This paper investigates the determinants of productivity for forty regions of Japan. We look at traditional determinants, such as Own-Size and Market Potential, as well as determinants more strongly linked to the recent literature on economic geography, such as forward and backward linkages. We also consider influences that have figured prominently in recent work, such as the MAR versus Jacobs debate on the role of regional diversity of production, and the role of landlocked status in productivity.

The most robust relations to productivity come from the Own-Size and Backward Linkage variables. Both the MAR externality and Market Potential variables are significant and the correct sign in the absence of the Backward Linkage variable. However they become insignificant or take on the wrong sign when it is included. While one can posit theories under which Forward Linkages may have a role in productivity, we do not find this in the data. Neither do we find that there is a productivity loss for regions of Japan that are landlocked.

Our estimates suggest an important link between region size and productivity. *Ceteris paribus*, a doubling of region size raises productivity by 3.5 percent. Backward linkages are also quite economically significant in accounting for differences across regions in productivity. A

³ We also intend to perform a counterfactual in which we preserve the current size distribution of regions but ban trade

simple counterfactual, premised on aggregate activity being spread evenly across the regions of Japan, would lower output by nearly 20 percent.

Taken together, these results suggest that there are quite important direct productivity gains associated with the concentration of economic activity in Japan. We must caution, though, that while we can quantify directly the productivity gains, a full consideration of welfare effects would likewise need to quantify costs arising from congestion, which falls beyond the scope of this paper.

References (Incomplete)

- Bernstein, Jeffrey and David Weinstein (1998) "Do Endowments Predict the Location of Production? Evidence from National and International Data," NBER #6815, University of Michigan.
- Davis, Donald R. and David E. Weinstein, (1996) "Does Economic Geography Matter for International Specialization," NBER # 5706, August 1996.
- Davis, Donald R., David E. Weinstein, Scott C. Bradford and Kazushige Shimpo (1997) "Using International and Japanese Regional Data to Determine When the Factor Abundance Theory of Trade Works," *American Economic Review*, June.
- Dixit, Avinash K., and Stiglitz, Joseph E. (1977) "Monopolistic Competition and Optimum Product Diversity," *American Economic Review* 67 (3), 297-308.
- Ellison,-Glenn; Glaeser,-Edward-L, "Geographic Concentration in U.S. Manufacturing Industries: A Dartboard Approach," *Journal-of-Political-Economy*; 105(5), October 1997, pages 889-927.
- Engel, Charles and Rogers, John H. (1996) "How Wide is the Border?" *American Economic Review*, December.
- Ethier, W. (1979) "Internationally Decreasing Costs and World Trade," *Journal of International Economics*, 9: 1-24.
- Fujita, M., Krugman, P. and Venables, A. (1999) *The Spatial Economy*, forthcoming MIT Press.
- Gallup, J.L, Sachs, J. and Mellinger, A. (1998) "Geography and Economic Development," mimeo, Harvard University.
- Glaeser,-Edward-L. et-al., "Growth in Cities," *Journal-of-Political-Economy*; 100(6), December 1992, pages 1126-52.
- Henderson,-J.-Vernon, "Efficiency of Resource Usage and City Size," *Journal-of-Urban-Economics*; 19(1), January 1986, pages 47-70.
- Henderson,-Vernon; Kuncoro,-Ari; Turner,-Matt, "Industrial Development in Cities," *Journal-of-Political-Economy*; 103(5), October 1995, pages 1067-90.
- Justman,-Moshe, "The Effect of Local Demand on Industry Location," *Review-of-Economics-and-Statistics*; 76(4), November 1994, pages 742-53.

Krugman, Paul R. (1980) "Scale Economies, Product Differentiation, and the Pattern of Trade," *American Economic Review*, 70, 950-959.

Krugman, Paul R. (1991) *Geography and Trade*, Cambridge: MIT.

Krugman and Venables (1995) "Globalization and the Inequality of Nations," *Quarterly Journal of Economics*, CX: 4.

Nakamura, Ryohei, "Agglomeration Economies in Urban Manufacturing Industries: A Case of Japanese Cities," *Journal-of-Urban-Economics*; 17(1), January 1985, pages 108-24.

Sveikauskas, Leo A., "The Productivity of Cities," *Quarterly-Journal-of-Economics*; 89(3), Aug. 1975, pages 393-413.

Sveikauskas, Leo; Gowdy, John; Funk, Michael, "Urban Productivity: City Size or Industry Size?" *Journal-of-Regional-Science*; 28(2), May 1988, pages 185-202.

Figure 1
Zipf's Law

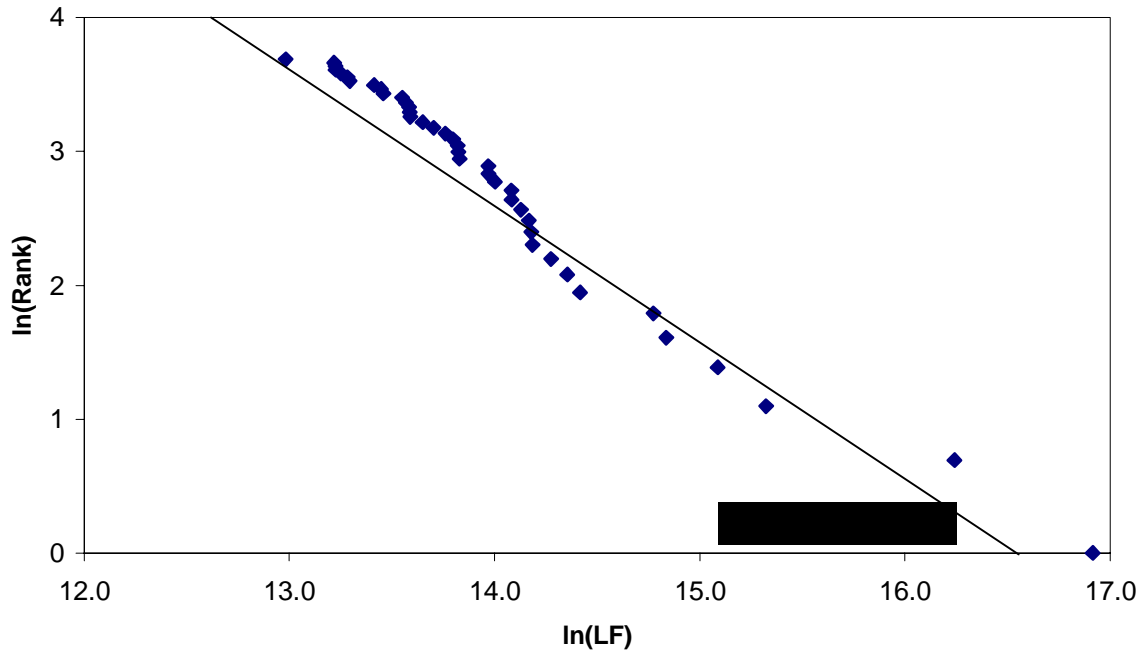


Figure 2
Productivity and Home Market Size

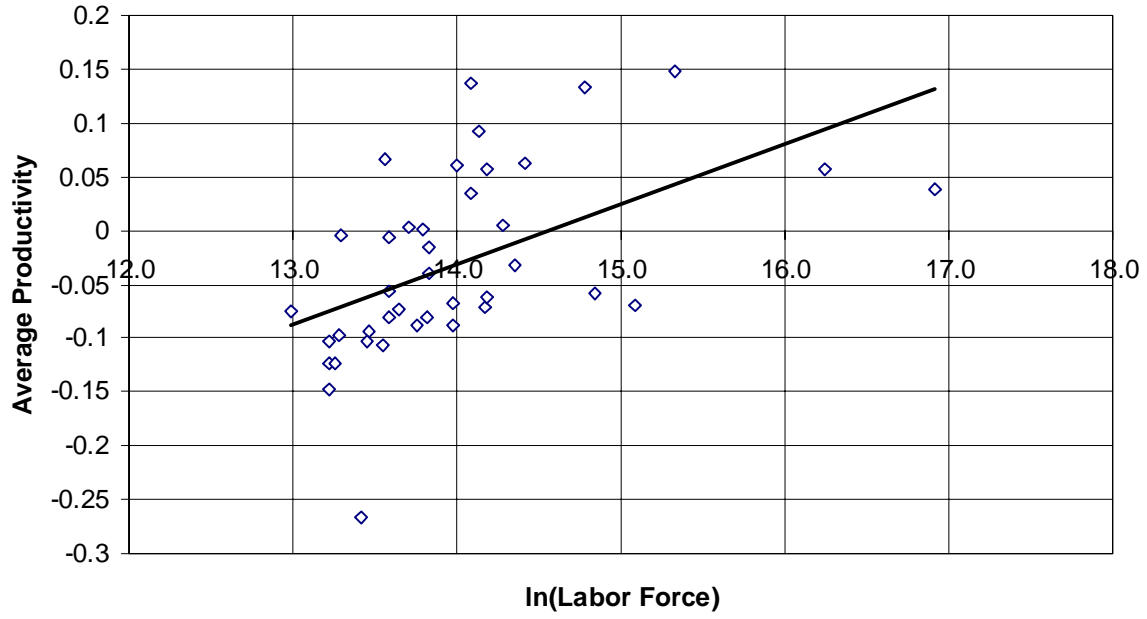


Figure 3

Productivity and Backward Linkages

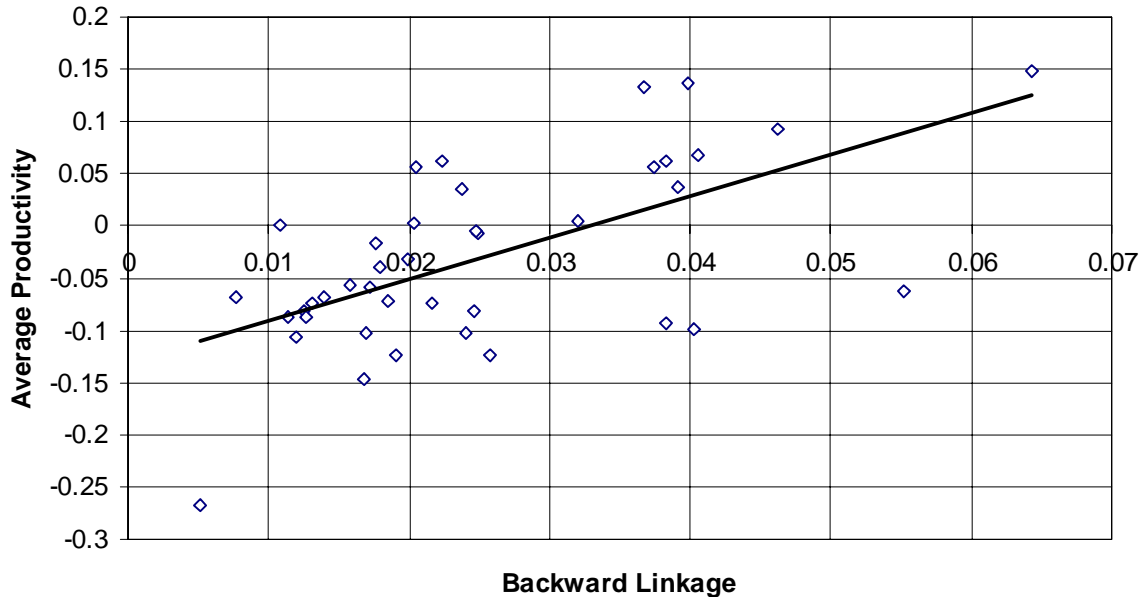


Table 1

Sample Statistics

Variable	Mean	Standard Deviation	Minimum	Maximum
Prod. of Non-College	-0.128	0.134	-0.341	0.206
Productivity of College	0.055	0.137	-0.298	0.379
Productivity of Capital	-0.020	0.086	-0.172	0.164
Market Potential	0.025	0.011	0.007	0.049
Forward Linkage	0.025	0.010	0.009	0.051
Backward Linkage	0.025	0.013	0.005	0.064
ln(Labor Force)	14.001	0.800	12.983	16.916
Specialization	2.024	1.059	0.828	5.390
Landlocked	0.150	0.362	0	1

Correlation Matrix

	NON	COLL	CAP	MP	FOR	BACK	ln(LF)	SPEC
Prod. of Non-College	1.000							
Productivity of College	0.214	1.000						
Productivity of Capital	0.374	0.285	1.000					
Market Potential	0.607	0.167	0.324	1.000				
Forward Linkage	0.648	0.170	0.377	0.970	1.000			
Backward Linkage	0.658	0.261	0.395	0.956	0.976	1.000		
ln(Labor Force)	0.679	0.155	0.251	0.328	0.358	0.352	1.000	
Specialization	0.089	0.125	0.174	0.399	0.308	0.327	-0.230	1.000
Landlocked	0.265	0.242	0.012	0.524	0.495	0.557	-0.043	0.211

Table 2

Determinants of Regional Productivity

	1	2	3	4	5	6
ln(Labor Force)	0.043 (0.007)		0.034 (0.006)		0.0234 (0.008)	0.041 (0.007)
Backward Linkage		3.956 (0.543)	3.716 (0.558)			5.992 (1.344)
Market Potential				5.097 (0.705)	4.360 (0.922)	-3.510 (1.923)
Log Likelihood	-773.804	-764.173	-759.506	-845.741	-765.263	-758.094
N	120	120	120	120	120	120

Dependent variable is regional factor productivity. Standard errors below estimates.

Table 3**Determinants of Regional Productivity:
Robustness Check of Alternative Explanations**

	1	2	3	4
ln(Labor Force)	0.056 (0.008)	0.037 (0.008)	0.031 (0.007)	0.046 (0.011)
Market Potential				-3.992 (2.021)
Backward Linkage		3.548 (0.620)	3.954 (0.620)	6.064 (1.321)
Specialization	0.032 (0.012)	0.006 (0.010)		0.013 (0.010)
Landlocked			-0.026 (0.028)	-0.011 (0.028)
Log Likelihood	-770.827	-759.324	-759.098	-757.317
Nobs.	120	120	120	120

Dependent variable is regional factor productivity. Standard errors below estimates.

Table 4

Determinants of Regional Productivity: Robustness Check
Using GDP Instead of Labor Force to Measure Region Size

	1	2	3
ln(GDP)	0.044 (0.006)	0.033 (0.005)	0.039 (0.007)
Market Potential			-3.222 (1.754)
Backward Linkage		3.662 (0.522)	5.636 (1.160)
Log Likelihood	-852.009	-836.589	-835.185
N	120	120	120

Dependent variable is regional factor productivity. Standard errors below estimates.

Table 5

Determinants of Regional Productivity: Robustness Check
Using Forward Linkages as well as Backward Linkages

	1	2	3	4
ln(Labor Force)	0.034 (0.006)		0.0313 (0.007)	0.037 (0.008)
Backward Linkage	3.716 (0.558)			6.442 (2.663)
Forward Linkage		5.547 (0.725)	5.111 (0.840)	-4.012 (3.833)
Log Likelihood	-759.506	-764.173	-761.314	-759.066
N	120	120	120	120

Dependent variable is regional factor productivity. Standard errors below estimates.