Identifying spatial efficiency–equity tradeoffs in territorial development policies: Evidence from Uganda*

Somik V. Lall, Elizabeth Schroeder**, Emily Schmidt

Abstract: We contribute to the debate on the spatial allocation of infrastructure investments by examining where these investments generate the highest economic return ("spatial efficiency"), and identifying tradeoffs when infrastructure coverage is made more equitable across regions ("spatial equity"). We estimate models of firm location choice in Uganda, drawing on insights from the new economic geography literature. The main findings show that manufacturing firms gain from being in areas that offer a diverse mix of economic activities. Public infrastructure investments in other locations are likely to attract fewer private investors, and will pose a spatial efficiency-equity tradeoff.

---

*Lall is with the World Bank, Schroeder is at Oregon State University and Schmidt is with IFPRI. Research for this paper has been funded by a DEC research grant , "Understanding the spatial efficiency – equity tradeoff s in regional development policies." We have benefited from discussions with Cecilia Briceno-Garmendia, Antonio Estache, Marianne Fay, Indermit Gill, Praveen Kumar, Dino Merotto, Marisela Montoliu, Christopher Timmins, Hyoung Wang and participants of a SDN Economists seminar on the efficiency and equity implications of alternate regional development policies. Data and codes available on request.
**Corresponding author: Department of Economics, 303 Ballard Extension Hall, Oregon State University, Corvallis, OR, 97331-3612. Phone: 541-737-1476. Fax: 541-737-5917.
Email: liz.schroeder@oregonstate.edu.

IDENTIFYING SPATIAL EFFICIENCY–EQUITY TRADEOFFS IN TERRITORIAL DEVELOPMENT POLICIES: EVIDENCE FROM UGANDA

Schroeder, Elizabeth; Lall, Somik; Schmidt, Emily

1. Introduction

Greater concentration of economic activity in a few places is part of the spatial transformation that accompanies development. The main message of the 2009 World Development Report *Reshaping Economic Geography* (World Bank 2008) is that economic growth will be geographically unbalanced, and trying to spread out economic activity is tantamount to discouraging it. On the other hand, living standards can spatially converge if policies facilitate economic integration between lagging and leading places. World Bank (2008) identifies that investments to improve health, education and information in lagging areas, along with efforts to encourage labor mobility, are most effective for economic integration.

There has been a great deal of debate in recent years over spatial industrial policy. For example, Special Economic Zones (SEZs) have become an increasingly popular policy tool to promote growth and development. Collier and Page (2009) argue that SEZs can be used in African countries to harness market forces and take advantage of agglomeration externalities in manufacturing clusters, allowing industries to break into global markets. Many policymakers, however, do not view increasing economic concentration as a beacon of progress. Rather, they actively try to stimulate economic growth in areas not favored by the market in an attempt to balance economic activity across the national territory. Are these policies fighting market forces of economic concentration? Or are they adding net value to the national economy by tapping underexploited resources?

In many countries, policymakers view the challenge of reducing economic distance across regions as one of reducing physical distance. Place-specific investments such as infrastructure, intended to reduce transport costs and improve accessibility of peripheral areas, are viewed as integral components of territorial development policies. Isolation from markets can reduce consumer welfare, as residents face higher prices due to market fragmentation and have less access to transport-dependent services. Isolation can also make it impractical for local producers to increase their scale of production and invest in cost-reducing technologies. Infrastructure investments that connect peripheral areas to markets should therefore improve consumer welfare and productive efficiency. This is particularly relevant to Uganda, where high transport costs are the most obvious 'natural barriers' that increase the costs of trade. Internal transport costs make exporting more costly, equivalent to an effective taxation rate of 22% for Ugandan exports on average in 1994 (Milner *et al*, 2000: 84). Infrastructure investments that reduce transport costs can deliver significant benefits.

There is little clarity, however, on whether spatial equity in transport coverage will lead to spatial equity in economic outcomes across places. Recent analytic work in economic geography points out that for activities that benefit from increasing returns to scale or agglomeration externalities, a fall in transport cost is unlikely to result in relocation of industry or growth of industry in lagging areas. In addition, there are concerns about whether spatial equity priorities tilt resource allocation away from regions where the returns will be highest and, by being "too" redistributive, take resources away from potentially high-return regions and impose significant efficiency costs at the national level.

Identifying such spatial efficiency-equity tradeoffs is at the core of designing territorial development strategies. These tradeoffs are rarely assessed, often because information on regional constraints to growth is limited and policymakers lack empirical evidence to inform their decisions. This paper contributes to the debate on the spatial allocation of infrastructure investments by examining *where* infrastructure investments will generate the highest economic returns (spatial efficiency) and identifying whether there are tradeoffs when infrastructure coverage is made more equitable across regions (spatial equity).

To identify spatial efficiency-equity tradeoffs, we examine the factors that entrepreneurs value when deciding where to locate production facilities, and how these decisions are influenced by improvements in infrastructure linking specific regions to market centers. Our empirical strategy is based on estimating models of firm location choice, with analytic underpinnings based on the new economic geography literature that develops linkages between infrastructure expansion, industrial clustering and agglomeration externalities (Fujita et al 1999; Puga 2002; Baldwin et al 2005) If firms value scale economies from market access and externalities from agglomeration, then they are likely to concentrate production facilities. Infrastructure investments in these places can relieve congestion costs and attract further private investment. However, infrastructure investments may also try to promote spatial economic equity by improving access to remote areas. These investments may be unsuccessful if the benefits they generate cannot offset the benefits that firms get from agglomeration externalities.

Thus, investments in remote areas may come at an opportunity cost to existing firms in dynamic areas, thereby creating a trade-off with national economic growth. Models developed in Baldwin et al (2005) analytically show that, in the presence of agglomeration effects, mobile firms are "locked in to" existing locations (and new ones are attracted to the same areas), thereby creating inertia in how firms respond to policies aimed at inducing relocation. The effects of infrastructure policies are likely to be insignificant until a threshold is crossed where the gains from relocation are higher than from staying.

Our identification of investment priorities focuses on two specific questions: (a) How much do infrastructure endowments matter when entrepreneurs make decisions on where to set up business establishments? (b) In comparison, do firms care about being physically close to other firms in the same line of business, or about locating in diverse economic environments? Once we know the relative valuation of infrastructure improvements and agglomeration economies for specific activities, we can identify returns to public investments in different locations.

Our empirical analysis focuses on Uganda for two main reasons. First, the country's national economic policy calls for infrastructure improvements to accelerate national economic growth as well as develop a regionally-balanced industrial landscape. Specifically, the National Industrial Policy identifies serious infrastructure shortfalls – particularly in electricity supply and transport - as being binding constraints to growth. It also makes a case for infrastructure to support industrial parks in 21 towns throughout the country to create a national portfolio of industrial centers (Government of Uganda, GoU 2007). Identifying the implications of public investments in stimulating private investment in alternative locations can improve the sharpness of the infrastructure portfolio.

Second, the Ugandan National Business Registry contains spatially-referenced data, providing detailed information on the location and product lines of industrial firms in the country. The physical location of these firms is identified with considerable accuracy using GPS technologies. By combining economic analysis with geographically referenced data on placement of infrastructure (roads and electric grids), natural topography, as well as distribution of human capital across the country, we can concretely identify locations which generate the highest economic returns to public infrastructure investments.

Our main findings highlight that establishments in the manufacturing industry gain from being in areas that offer a diverse mix of economic activities. The economic geography literature points out that economic diversity, which is synonymous with urbanization externalities, is associated with increased access to a broad range of producer and consumer services such as business, legal, and financial services. Typically, economic diversity increases with size of the agglomeration. In addition, availability of power supply, transport links connecting districts to markets, and the supply of skilled workers attract manufacturing activities. Combining all these factors gives a distinct advantage to existing agglomerations. In Uganda, this means urban areas around Kampala and Jinja are likely to lead Uganda's industrial development. Infrastructure investments that improve conditions for growth in these areas are likely to produce the highest returns compared to investments elsewhere. From a spatial efficiency perspective, these should be high-priority public infrastructure investment locations as Ugandan policymakers consider policies for accelerating growth. Public infrastructure investments in other locations are likely to attract few private investors, and are likely to pose an economic efficiency-equity tradeoff. More distant areas may benefit locally through basic improvements in connectivity by easing their access to downstream industry and markets; equity considerations may suggest a basic level of investment in these areas, particularly for primary producers who face location constraints. The aggregate economic returns in these areas, however, are lower than such investments being placed around urban conurbations.

Following this introduction, Section 2 reviews the literature, Section 3 specifies the estimation strategy, and Section 4 describes the data and clustering of manufacturing. Section 5 discusses the findings from the empirical analysis and provides alternate scenarios for transport improvements, and Section 6 concludes.

2. **Literature**

The clustering of manufacturing establishments raises an important policy question: are firms clustering because they are constrained in their location choice by a need to be near sparse infrastructure networks, or are there gains from exploiting agglomeration economies? If it is the former, then policies such as transport and electricity network expansion (which lower transport and production costs in the peripheral regions) can allow firms to move to lower cost locations and still access markets. If agglomeration economies dominate, then it may be useful to improve infrastructure services in congested areas to maximize positive spillovers associated with industrial development.

In making decisions about where to set up businesses, entrepreneurs are most likely to select areas that offer conditions where profits can be maximized. Prices and quality of inputs, prices of outputs and access to technology matter. Firms are likely to cluster in areas that provide good access to markets, as the size of the market influences the firm's decision to increase scale and invest in cost-reducing technologies. Firms may also be attracted to areas that already have firms established in their lines of business due to localization economies; new firms can learn from existing ones about business processes, new technologies and informal regulations, as well as benefit from a pool of trained workers. Finally, firms may value the overall economic diversity of an area. These are often referred to as urbanization economies, and are associated with good access to a broad range of producer and consumer goods that typically increase with size of the agglomeration.

A recent paper surveying industrial location decisions in developing countries identifies the following factors as being important (Deichmann et al 2008):

- Factor prices.
- Quality and cost of complementary utility services, including electricity, water and telecommunication.
- Market access as a function of the size of the region that can be reached given existing transport infrastructure.
- Agglomeration economies as measured by the presence of firms in own industry and of firms in related, e.g., buying or supplying, industries.
- Labor and other regulations.

In many of the papers covered in the survey, benefits of agglomeration economies (both own-industry and overall diversity), market access and infrastructure endowments outweigh the costs imposed by congestion, increasing wages and land prices (Deichmann et al 2008). Using firm survey data from India, Lall and Mengistae (2005) find that localization economies, as measured by own-industry concentration, have significant bearing on firm location decisions across cities. This effect is the highest for technology-intensive sectors. Deichmann et al (2005) find similar evidence for manufacturing firms in Indonesia. Here, localization effects are more important for high-technology (e.g., office computing) and natural resource-based industries (such as wood or rubber and plastic).

Empirical work on urbanization economies is mixed. Empirical studies for the United States show that diversity in economic activity has positive impacts on regional economic growth (Bostic 1997; Garcia-Mila and McGuire 1993; Glaeser et al 1992). On the other hand, also using data for the United States, Mirachy (1995) finds little evidence to support the diversity argument. For India, Lall et al (2003) find evidence that diversity is the most important source of external cost reduction for Indian manufacturing establishments. Their analysis is based on estimating cost functions with micro data for specific manufacturing industries.

For Indonesia, Henderson et al (1995) show that the relative importance of urbanization economies is higher in new high-tech industries compared to mature capital goods industries. These findings are consistent with product cycle theory (Vernon 1966) and insights from work on “nursery cities” (Duranton and Puga 2001), which predict that new industries tend to prosper in large and diverse urban areas, but with maturity, their production facilities move to smaller and less diverse cities.

Using a unique spatially-referenced dataset of Ugandan manufacturing firms, we estimate a location-choice model to understand the main factors that influence decisions of entrepreneurs to establish manufacturing establishments across areas of the country. As discussed below, we know that manufacturing overall is clustered. We want to know if this clustering is due to benefits from localization economies, or driven by transport links that connect areas to markets, availability of complementary production inputs such as electricity, the quality of the local labor force, or the benefits from being in a diverse economic environment. Ugandan manufacturing is not technology-intensive or innovation-led. It is dominated by production activities that are standardized and require low technology by global standards. These include food processing, garments and textiles, clay products and furniture; access to domestic markets for raw materials is therefore important, as well as access to forward linkages. However, many of these products and business lines are new to the country, so they can be considered locally as “sunrise” activities, while being “sunset” activities globally.

3. Methodology

To examine the location decisions of firms, we specify a profit function in which an establishment will be located in a particular region if the profits from being there are higher than profits in any other region of the country. This model is an adaptation of the Bayer and Timmins (2007) equilibrium model of location choice to the question of industrial development. In the model, profits π earned by establishment *i*, in industry *k*, which chooses to locate in region *j* are:

$$\pi_{i,j,k} = f\left(\sigma_{j,k}, A_j, IR_j\,, LIN_j, H_j, X_j, \eta_{i,j,k}; \bar{\beta}_k\right) \qquad (1)$$

Agglomeration effects that provide production externalities are represented by $\sigma_{j,k}$ (localization economies) measured as the own-industry concentration of industry *k* in region *j*; A_j represents externalities from urbanization economies (measured by industry diversity); IR_j refers to the quality and availability of inter-regional infrastructure that links the region to market centers; LIN_j reflects local infrastructure conditions in the region, such as power supply; H_j represents the region-specific stock of human capital. In addition, X_j refers to region-specific natural geography conditions, which include ruggedness of the region’s terrain, natural resources to support development, and climate (rainfall). Good natural geography (“first advantage” in the expression used by Burgess and Venables (2004)) is likely to stimulate early period population growth and economic development, and neglecting these factors could provide misleading effects of the economic geography variables (market access and agglomeration economies).

We choose the following functional form for the profit function:

$$\pi_{i,j,k} = \beta_1\sigma_{j,k} + \beta_2 A_j + \beta_3 IR_j + \beta_4 LIN_j + \beta_5 H_j + \beta_6 X_j + \eta_{i,j,k} \qquad (2)$$

We estimate a set of coefficients, β, for the full sample and separately for each industry *k*. The i^{th} firm will choose region *j* if $\pi_{i,j,k} \geq \pi_{i,l,k}$ for all *l*, where *l* indexes all the possible region choices to the i^{th} firm. For estimation, we will assume that $\eta_{i,j,k}$ is additively separable from the rest of the utility function and has a Weibull distribution. The result is that we can write the probability that any firm will choose to locate in region *j* (McFadden 1973)

$$P(\pi_{i,j,k} \geq \pi_{i,l,k}, \forall\, l \neq j) = \frac{e^{\beta\sigma_{j,k}+\beta A_j+\beta IR_j+\beta LIN_j+\beta H_j+\beta X_j}}{\sum_{l=1}^{j} e^{\beta\sigma_{l,k}+\beta A_l+\beta IR_l+\beta LIN_l+\beta H_l+\beta X_l}} \qquad (3)$$

In our estimation, we are assuming that each firm takes attributes associated with each region as given and makes rational location-choice decisions. For the purpose of estimation, this assumption translates into a condition in which the idiosyncratic error term is independent of the regional characteristics. One of the main empirical challenges to separately identifying the effects of local spillovers (i.e. localization economies) is that the concentration of firms in location *j* may be correlated to sources of natural advantage that are not observed in the data. If favorable natural conditions encouraged or facilitated concentration of firms in particular areas, then not addressing this correlation is likely to overstate the impact of agglomeration economies.

A standard solution for this omitted variable problem would be to employ instrumental variables. However, we use a conditional logit model to estimate equation (3), which implies that we cannot use instrumental variables. To address this problem within our estimation framework, we only analyze location decisions of firms that have started business in the four years preceding the survey. Next, we create agglomeration variables using data for establishments that were in business five or more years before the survey. By splitting the data, we hope that unobserved characteristics that matter for today's location decisions are different from those that influenced previous concentration. Our second strategy is to include a range of variables representing sources of "first advantage" directly into the estimation. These variables, described above, should capture why some areas became attractive for people and establishments in the first place.

4. Data, Stylized Facts and Variable Construction

Clustering of Industrial Activity

Much of the industrial activity in Uganda is clustered around large cities and along transport corridors (Figure 1). Mapping the location of industrial firms onto the country's geographic profile makes it clear that industrial activity in Uganda is concentrated (see the map on the left in Figure 1). Most of the country's 12,000 manufacturing firms with 5 or more employees are clustered along the industrial corridor stretching between the country's major urban agglomerations—Masaka, Kampala, Jinja, and Mbale. Also clear is that the location of industrial activity closely follows the distribution of infrastructure networks. Seventy percent of the manufacturing firms with 5 or more employees are located within 10 kilometers of a major road. In addition, most of these firms are located in regions that are close to national markets, measured using travel times to cities of 100,000 or more (see the map on the right of Figure 1).

Figure 1 about here

These manufacturing data are drawn from the Uganda Business Registry of 2001 (Uganda Bureau of Statistics, 2001), which provides a comprehensive listing of all establishments in the country. A total of 165,000 establishments are included in this database. For each establishment, we know its physical location (measured by a GPS system), its four-digit industrial classification, the year that it started operations, and the number of employees. The GPS coordinates provide each establishment with a unique location identifier (latitude and longitude) at accuracies of ten to fifteen meters. To our knowledge, the Uganda Business Registry is the only nationally representative database in developing countries that has identified establishments with such accuracy.

Using this database, we examine the extent to which manufacturing activity in Uganda is *localized* – i.e. concentrated across locations. Our data make it possible to examine distances between firms directly. In contrast, most establishment-level datasets force researchers to analyze the data using administrative units as the units of observation, so often valuable information on localization is lost due to aggregation. Not surprisingly, commonly used measures of localization have been implemented with aggregate data in mind. Such examples are the Ellison-Glaeser (Ellison and Glaeser, 1997) and location quotient (Isard, 1956) indices.

To exploit the unique feature of our dataset, we calculate establishment-to-establishment distances for all manufacturing firms and for specific industries. This is similar to the procedure employed by Duranton and Overman (2005), in which they set up a measure to examine localization using micro geographic data. We start by calculating the Euclidian distance between every pair of manufacturing establishments. For manufacturing industry *M* with *n* establishments this generates $n(n-1)/2$ unique distances between establishments. We then calculate the frequency for each distance level and plot the corresponding density. We can represent the Euclidian distance between establishments *i* and *j* by $D(i,j)$, and define $\delta(i,j,d)$ such that $\delta(i,j,d) = 1$ when $D(i,j) = d$ and $\delta(i,j,d) = 0$ otherwise. The un-smoothed distance density or *K*-density is:

$$K_M(d) = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \frac{\delta(i,j,d)}{n(n-1)}$$

For overall manufacturing, we find small distances between establishments. A third of all manufacturing firms are within 3.7 kilometers of each other; 45 percent are within 9 kilometers, and the 50th percentile is 41.5 kilometers. The distribution of overall distances is shown in Figure 2. The distribution shows a kernel density of distances, which is plotted using the Gaussian kernel specification in STATA.

Figures 2 & 3 about here

We also looked at specific industries to examine the extent to which firms were clustered. For instance, for establishments in paper and printing (SIC 222), the 50^{th} percentile of inter-establishment distance is 1.8 kilometers, signifying that the industry is much more localized relative to manufacturing as a whole. In comparison, the 50^{th} percentile of inter-establishment distances for garments (SIC 181) is 9 kilometers, and 89 kilometers for food processing (SIC 153). The distributions of inter-establishment distances are show in Figure 3. The descriptive statistics in this section have shown that manufacturing activity is clustered.

Variable construction

We here outline the variables employed; technical details on data sources (Thomas, 2007a and 2007b; Uganda Bureau of Statistics, 2001, 2002a, 2002b and 2006) and the estimation methods are in the online appendix.

Infrastructure for market access: We computed travel times on the road network from each firm to the nearest city of 100,000 or more people. Estimation of travel times are presented in the online appendix.

Power supply: A dummy variable showing the presence of an electric grid in the district is used to proxy for access to power supply. This is a crude measure, but we could not obtain data on actual usage or reliability of power supply.

Localization: There are several ways of measuring localization. These include own-industry employment in the region, own-industry establishments in the region, or an index of concentration, which reflects disproportionately high concentration of the industry in the region in comparison to the nation. We use own-industry establishments as the measure of choice, as we want to test whether there are gains from locating in areas that are already specialized in the firm's chosen line of business. Also, when location "shopping," entrepreneurs are more likely to observe the density of establishments in an area, compared to the number of people employed in them. Using the number of firms in the same sector has also been common in empirical work, with the underlying premise that localization economies come from the absolute volume of similar activities in the neighborhood.

For the localization measure, we consider the number of establishments in the same 2-digit industry sector that are within 20 kilometers of each firm under consideration. As discussed earlier, we only use establishments that have been in business for more than 5 years for this calculation. For the location modeling, we then average these numbers at the district level.

Economic diversity (urbanization economies): We use a region's economic diversity to reflect potential gains from urbanization economies. Typically, larger cities have a greater diversity of firms (Deichmann et al. 2008). This allows greater specialization since it enables small, innovative firms to access a larger pool of potential buyers and complementary services that cannot be provided in-house. Larger cities also provide a larger home market for end products, make it easier to attract skilled employees who are attracted by urban amenities not available in smaller towns, and support a large number of complementary service providers such as financial and legal advisers, advertising and real estate services.

The well-known Herfindahl measure is used to examine the degree of economic diversity in each district. The Herfindahl index of a district *j*, H_j, is the sum of squares of employment shares of all industries in district *j*:

$$H_j = \sum_k \left(\frac{E_{kj}}{E_j}\right)^2 \quad (4)$$

Unlike measures of specialization, which focus on one industry, the diversity index considers the industry mix of the entire regional economy. The largest value for H_j is 1 when the entire regional economy is dominated by a single industry. Thus a higher value signifies a lower level of economic diversity. For more intuitive interpretation of the measure, therefore, the diversity index in our model is H_j subtracted from unity: $DV_j=1-H_j$. A higher value of DV_j signifies that the regional economy is relatively more diversified.

Variables on district-specific characteristics were computed from various sources. The human capital variable reflects the share of each district's working age population with primary school or higher education. This comes from the 2002 Uganda Census. Other variables such as terrain roughness were derived from USGS/NASA SRTM data. Computation of roughness per district is described in the online appendix. In addition, we were also able to disaggregate major crop production by district using a Spatial Allocation Model (SPAM) developed by IFPRI (see You et al., 2007 and You, Wood and Wood-Sichra, 2007). The complete list of data sources is provided in the online appendix.

To address concerns about collinearity in the independent variables, we compute the variance inflation factor (VIF) for every variable in each of the regressions. A rule of thumb is to be concerned when the VIF exceeds 10. In our model, the VIF of the education variable exceeds 10 in every specification. Indeed, the education variable is found to have a correlation of 0.593 with the localization variable and -0.756 with the market access variable. Since education is significant at a 5% or 1% in every regression other than one industry-specific specification (Chemicals and Petroleum), however, the education VIF is not of great concern. Only one other variable, presence of an electric grid, has a VIF at or near 10 in any specification. As discussed above, this is a rather crude measure of access to a power supply, but its high VIF could be a reason that it loses significance in some of the industry-specific regressions.

5. Results and Discussion

The sample for estimating the location choice model includes all firms in the Uganda Business Registry that have more than five employees and that were less than five years old at the beginning of the survey. We limit the sample to relatively new entrants because older firms may have made location decisions facing considerably different location attribute choices. There are 56 districts (using 2002 definitions) that firms can chose between in Uganda. The number of new-entrant firms in the sample is 1,603, resulting in 89,768 observations on potential firm locations. In general, the model performs very well in predicting where establishments will be located. Based on the model parameters, our success in predicting actual location decisions is 98 per cent. Table 1 provides the raw estimates and standard errors from the conditional logit model. Column 1 reports estimates for all manufacturing firms. Columns 2-8 provide sector-specific estimates. These sectors are: Food and beverages, textiles and apparel, paper and printing, chemicals and petroleum, rubber and plastics, metal products and furniture.

Table 1 about here

<u>Infrastructure</u>: We find that access to the power grid has a positive effect on a district's attractiveness for location of manufacturing activity. These results are significant in estimations for all manufacturing. While the estimates are positive for each of the industry sectors, they are statistically significant for food and beverages, garments and textiles, and furniture industries. Keep in mind that our measure of power supply is a crude one – we only have information on whether or not the power grid runs through the district. It would be useful to collect information on power breakdowns and prices for future analysis. These findings are similar to those obtained in analysis of location decisions of Indian manufacturing (Lall and Mengistae 2005, Mani et. al 1997).

Market access, measured by transport connectivity to cities of 100,000 or more people, is an important factor in determining industry location. Remoteness from market centers lowers industrial prospects. Pooled estimates for all manufacturing industries produce statistically significant effects. For specific industries, establishments in food and beverages, as well as and chemicals and petroleum products, value market access (after controlling for the other variables). Estimates for other sectors are not statistically significant.

Agglomeration: Given the extent of clustering seen in the data, one would expect that the presence of own-industry concentration would directly influence location choices. This is correct when we use the localization variable as the only determinant of industry location. However when we control for other factors, the localization variable has a negative effect on location decisions in models using all manufacturing establishments in the estimation. This would imply that competition and prices of fixed production factors increase with industry agglomeration – and would make clustered locations more expensive. However, results for individual sectors exhibit considerable heterogeneity. The effects of localization are positive and significant for establishments in food and beverages, chemicals, rubber and plastics, metals and the furniture industry. However, for the paper and printing industries, localization economies have a negative effect on location choices. Given these mixed signals from localization industries, why do establishments concentrate production facilities?

The answer to this puzzle is in positive economies that establishments accrue from economic diversity. The estimates of economic diversity for all manufacturing as well as specific industries are positive and significant. The only exception is chemical and petroleum products, where the estimate is not statistically significant. These results tell us that entrepreneurs locate establishments in areas that offer a diverse range of economic activities. In the economics literature, there are three main factors that explain the importance of economic diversity: (1) information sharing and innovation – large cities are breeding grounds for new ideas and innovations due to the concentration and diversity of knowledge sources. This facilitates product and process innovation, and therefore new products are more likely to be developed in diversified cities (Duranton and Puga 2001); (2) establishments located in large cities have relatively better access to producer amenities - such as business services, finance, logistics, banking, advertising, and legal services – which can enhance economic performance (Abdel-Rehman 1988, Fujita 1988, Rivera Batiz 1988); and (3) on the consumption side, increasing the range of local goods enhances welfare of households. Thus, economic diversity can yield external scale economies through the variety of consumer and producer goods.

Human capital: We find that the availability of workers who have primary or more schooling has a considerable impact on location decisions in manufacturing industries. The only exception is the chemical and petroleum industry, where the effects are not statistically significant. In general terms, a pool of semi-skilled workers makes it easier for firms to scale up production by hiring more workers. In fact, investment climate surveys in many developing countries identify the lack of skilled workers as a major impediment to increasing firm size and productivity.

Identifying high return areas

We now move from describing the empirical analysis of location choices to identifying *where* public infrastructure investments will produce the highest economic returns in terms of national industrial promotion. Addressing this policy concern requires that we recall how firms in various sectors value district-specific endowments. In particular, agglomeration economies from economic diversity are important location determinants. As firm performance depends on being located in a diverse and large urban environment, it is extremely difficult for policies to successfully move and sustain these activities in secondary locations. This is because successful relocation policies will need to coordinate decisions of firms across sectors. In addition, the stock of human capital is important for location decisions of manufacturing establishments.

At least in the short-to-medium term, these 'preconditions' are likely to be fixed, and the effects of infrastructure improvements will depend on the relative 'stock' of these attributes across districts. For example, consider educational attainment. In Kampala, 421,000 people of working age have completed primary or higher education. In comparison the 'human capital stock' in Lira is 92,000 and 56,700 in Gulu. Other things being equal, for industries that value skilled labor, Kampala becomes more attractive than upcountry centers.

Firms consider a package of amenities that a region offers in making location decisions. From our model, we can predict the relative profitability to manufacturing firms of locating across districts using the functional form in equation (2). Figure 4 plots these values, where each district's profitability is compared to that of Kampala, which is normalized to 100. From this figure, it is clear that expected profits are highest in Kampala, Wakiso and Jinja. These are high-return areas for private manufacturing. Districts at the borders of these agglomerations and those along the road leading to the Kenyan border also offer profitable opportunities for manufacturing. Table 2 lists the relative profits.

Figure 4 and Table 2 about here

In addition to overall profits, we can also calculate how specific attributes contribute to profitability differences. The effect of any particular attribute x_{jm} on the probability that a firm locates in district *j* is given by

$$\frac{\partial lnP_j}{\partial lnx_{j,m}} = \beta_m x_{j,m}(1 - P_j) \qquad (5)$$

where β_m is the parameter estimate from the clogit model, and P_j is the probability of firm location in district *j*. Consider, for example, the returns to a district of increasing human capital, given by the increase in percentage of firms choosing to locate there, for establishments in the food and beverages industry. These elasticities are mapped in Figure 5, and are relatively higher in the Kampala region.

Figure 5 about here

Geographically prioritizing infrastructure improvements

Given the distribution of relative profits and the 'preconditions' for success, where will infrastructure investments produce the highest returns? To examine this question, we simulate road improvements in two locations:

- In the first case, we simulate improvements of road conditions around the Northern cities of Gulu and Lira to increase travel speeds from 60 to 100 km/h. This would reduce the time it takes to travel from these cities to market centers of 100,000 people from an average of 5 hours to 3 hours.

- In the second case, we simulate improvements of roads around high profit cities– Iganga, Mpigi and Mubende, which are in the country's main industrial agglomeration. Again these road improvements are assumed to increase travel speeds from 60 to 100 km/h.

For eaxmple, in Iganga, a district that adjoins Kampala, transport improvements increase the share of establishments that would locate in the district from 5.8 per cent to 10.5 per cent. However, there are only small gains from additional investments in Mpigi and Mubende. Overall we find that the "pull" of agglomeration economies is strong and reduces the impact of complementary investments to decentralize manufacturing activity. This simple simulation exercise uses results from the empirical analysis and identifies that private returns to public infrastructure investments, measured by new industrial development, is highest in areas that offer 'preconditions' for success. In particular, the stock of human capital and an existing mix of diverse economic activities are important ingredients in a successful growth 'recipe'. Incidentally, these preconditions are offered in the country's main urban agglomerations.

Robustness

To examine the sensitivity of our results to modeling specifications, we ran several variations of the location choice model (see online appendix Tables A1-A5). First, we used different measures for our localization and market access variables. The localization measure, number of firms in the same industry within 20 kilometers of a given establishment, was replaced with the number of own-industry firms within 5 kilometers or 60 kilometers. In general, we found that agglomeration economies are higher at closer distances. We also replaced our market access variable with travel time to Kampala, which is by far the largest city in Uganda. This measure of market access was only significant for the food and beverage industry.

While conditional logit is the only model specification that allows us to identify the impact of the characteristics of each district on the probability of a firm locating there, we estimated different variations of the conditional logit. First, in the all-manufacturing sample, we included interactions between the district characteristics and a set of dummy variables for sector. To further explore the impact of sector, we allowed the sector dummies to directly impact the probability of choosing each district by estimating what is sometimes known as a "mixed" conditional logit, which allows for variables that vary only by firm. Neither of these is our preferred specification, because the proliferation of coefficients to be estimated leads to a loss of power, particularly in the mixed logit, where none of the sector coefficients is precisely estimated. These results are consistent with our initial findings, however, with the market access, electric grid, diversity index, and education variables all remaining significant at similar magnitudes to those of the primary specification. We also tried varying the choice sets of the firms, to the extent allowed by our data. When the location choice is among one of Uganda's four regions, rather than districts, there are no contradictions in significant results. Similarly, when the choice is among fourteen sub-regions, our main findings are confirmed – access to markets and a diverse local economy are significant drivers of firm profits and thus location decisions. In both specifications, however, we lose some power.

Finally, we controlled for additional physical characteristics of each district by adding controls for agricultural production. Controlling for the tons of coffee, cotton and maize produced in each district did not change any of our results qualitatively, although again some power was lost. The results of these robustness checks are available in the online appendix.

5. Conclusion

In this paper we find that entrepreneurs in Uganda value agglomeration economies, human capital, and infrastructure conditions in deciding where to locate manufacturing establishments. The effects of infrastructure improvements to promote industrial development and accelerate national economic performance are highest in areas that offer external scale economies from agglomeration and availability of skilled workers. These 'preconditions' are relatively abundant in the main urban agglomerations of the country – thus, improving infrastructure in these places provides the highest private return to public investment. On the other hand, using infrastructure to support economic growth in areas which are deficient in these 'preconditions' is likely to yield low returns. Investments to link peripheral regions to markets are also likely to be more expensive in absolute terms. Policymakers should consider these spatial efficiency-equity tradeoffs in deciding the spatial allocation of infrastructure investment.

The results from the location choice analysis are consistent with very detailed cost benefit analyses of transport improvement projects in Uganda (World Bank 2004). The World Bank's HDM model allows for the modeling, through time, of the interaction between traffic volume and composition, road condition and vehicle operating costs. The cost-benefit analysis using the HDM model shows that the net present value (NPV) of improving 67 km of roads between Kampala, Gayaza, Zirobwe and Wobulenzi, connecting the capital to agriculture-rich areas, was US$23.3 million. In comparison, improving 114 km of roads between the Northeastern towns of Soroti and Lira produced a NPV of US$9.9million.

At an annual average of US$21 per capita, current infrastructure spending in Uganda is extremely low given the current state of infrastructure services (Briceno-Garmendia, 2006). These levels are half the lowest annual average per-capita amount spent in Latin American countries at the end of the 90s, and they are only comparable to what Indonesia was spending on infrastructure immediately after the financial crisis. There is urgent need to scale up infrastructure investments.

However, infrastructure investment decisions need to be made in a way that responds to the country's development objectives at the lowest possible cost. If, as reflected in Uganda's National Industrial Policy (GoU 2007), industrial development is the cornerstone of the country's accelerated growth strategy, then infrastructure investments need to be prioritized towards modes and geographic areas that can produce the highest returns in terms of industrial development. The analysis in this paper provides one approach for prioritizing these investments.

REFERENCES

Abdel-Rahman, H. (1988). Product differentiation, monopolistic competition, and city size. *Regional Science and Urban Economics*, 18, 69-86.

Baldwin,R., R. Forslid, P. Martin, G. Ottaviano, and F. Robert-Nicoud (2005) *Economic Geography and Public Policy*. Princeton, NJ: Princeton University Press

Bayer, P and C. Timmins (2007) Estimating Equilibrium Models of Sorting Across Locations, *Economic Journal*, 117, 353-374.

Bostic, R. (1997). Urban productivity and factor growth in the late 19th century. *Journal of Urban Economics* 4, 38-55.

Briceno-Garmendia, C. (2006). Uganda - Infrastructure for Growth. Contribution to Uganda Country Economic Memorandum, Washington, DC: World Bank.

Burgess, R. and A. J. Venables. (2004). Towards a microeconomics of growth. In Bourguignon F. and B. Pleskovic (Eds.). *Annual World Bank Conference on Development Economics 2004: Accelerating Development* Washington, DC: World Bank.

Collier, P., and J. Page. (2009). *Industrial Development Report 2009.*Vienna: UNIDO

Deichmann, U, S.V. Lall, S. Redding, and A.J. Venables (2008). Industrial location in developing countries. *World Bank Research Observer*, 23

De la Fuente, A. And X. Vives (1995), Infrastructure and education as instruments of economic policy: Evidence from Spain, *Economic Policy*, 20

Duranton, G. and H. Overman (2005). Testing for Localization Using Micro-Geographic Data, *Review of Economic Studies*, 72. 1077-1106,

Duranton, G., and D. Puga (2001). Nursery Cities: Urban Diversity, Process Innovation, and the Life Cycle of Products, *American Economic Review*, 91, 1454-1477.

Ellison, G., & E. L. Glaeser (1997). Geographic concentration in U.S. manufacturing: a dartboard approach. *Journal of Political Economy*, 105, 889-927.

Estache, A. and M. Fay (2007). Current debates on infrastructure policy. Paper Prepared for the Growth Commission.

Fujita, M. (1988). A monopolistic competition model of spatial agglomeration: differentiated product approach. *Regional Science and Urban Economics*, 18, 87-124.

Fujita, M., P. Krugman, and A. J. Venables. (1999) *The Spatial Economy: Cities, Regions, and International Trade*. Cambridge, MA: The MIT Press.

Garcia-Mila, T., and T. McGuire (1993). Industrial mix as a factor in the growth and variability of state's economies. *Regional Science and Urban Economics* 23, 731-748.

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

Government of Uganda (2007). National Industrial Policy: A Strategy for Uganda's Transformation, Competitiveness and Prosperity. Kampala: Ministry Of Tourism, Trade & Industry (www.mtti.go.ug).

Henderson, J.V. (2003). The Urbanization Process and Economic Growth: The So-What Question, *Journal of Economic Growth*, 8, 47-71.

Isard, W. (1956). *Location and Space-Economy*. Cambridge, MA: MIT Press.

Jarvis A., H.I. Reuter, A. Nelson, E. Guevara. (2006). *Hole-Filled Seamless SRTM Data V3*. International Centre for Tropical Agriculture (CIAT). Available from http://srtm.csi.cgiar.org. Downloaded March 28, 2007.

Lall, S. V., J. Koo and S. Chakravorty (2003). Diversity Matters: The Economic Geography of Industry Location in India. *World Bank Policy Research Working Paper No. 3072*, Washington, DC: World Bank.

Lall, S.V. and T. Mengistae. (2005). The impact of business environment and economic geography on plant level productivity: an analysis of Indian industry. *Policy Research Working Paper 3664*. Washington, DC: World Bank.

Lall, S.V. and H. G. Wang (2007). Improving the Development Impact of Infrastructure – proposal for a research program. Mimeo, Washington, DC: World Bank.

Mani, M., S. Pargal and M. Huq (1997). Is there an environmental 'race to the bottom'? Evidence on the role of environmental regulation in plant location decisions in India. *Poverty, Environment and Growth Working Paper No. 15*, Washington, DC: World Bank.

Milner, C., O. Morrissey and N. Rudaheranwa (2000), 'Policy and non-Policy Barriers to Trade and Implicit Taxation of Exports in Uganda', *Journal of Development Studies*, 37, 67-90

Mirachy, W. (1995). Cities and the Product Cycle. Unpublished PhD Dissertation, MIT.

Puga, D. (2002). European regional policies in light of recent location theories. *Journal of Economic Geography*, 2, 373-406.

Republic of Uganda (2006). *Draft Estimates of Revenue and Expenditure (Recurrent and Development). 2005-2006*, Kampala: Ministry of Finance, Planning and Economic Development.

Rivera-Batiz, F. (1988). Increasing returns, monopolistic competition, and agglomeration economies in consumption and production. *Regional Science and Urban Economics*, 18, 125-154.

Thomas, T. S. (2007a). *Terrain Roughness, 10 km Aggregate, for Africa*. Washington, DC: The World Bank, FEU, SLD.

Thomas, T. S. (2007b). *Travel Time Friction Grid for Africa at 500-meter Resolution*. Washington, DC: The World Bank, FEU, SLD.

Uganda Bureau of Statistics (2001). Uganda Business Inquiry 2000/2001, Entebbe: UBOS.

Uganda Bureau of Statistics (2002a). Uganda Business Register 2001/2002, Entebbe: UBOS.

Uganda Bureau of Statistics (2002b). Population and Housing Census 2002, Entebbe: UBOS.

Uganda Bureau of Statistics (2006). Uganda National Household Survey 2005/2006, Entebbe: UBOS.

USGS Global GIS Database: Africa CD, VMAP level 0. 2001.

Vernon, R. (1966). International investment and international trade in the product cycle. *Quarterly Journal of Economics*, 80, 180-207.

World Bank (2004). Project in support of the third phase of the road development program. *Project Appraisal Document 27585-UG*, Washington, DC: World Bank.

World Bank (2008). *World Development Report 2009: Reshaping Economic Geography*. Washington, DC: Oxford University Press.

You, L and S. Wood (2003) Spatial allocation of agricultural production using a cross-entropy approach. *Environment and Production Technology Division Discussion Paper No. 126*, Washington, DC: International Food Policy Research Institute.

You, L., S. Wood, U. Wood-Sichra, J. Chamberlin. (2007). Generating plausible crop distribution maps for Sub-Sahara Africa using a spatial allocation model. *Information Development,* 23, 151-159

You, L., S. Wood, U. Wood-Sichra. (2007). Generating plausible crop distribution and performance maps for Sub-Saharan Africa using a spatially disaggregated data fusion and optimization approach. *IFPRI Discussion Paper 720*, Washington, DC: International Food Policy Research Institute.

Table 1: 'Raw' estimates from conditional logit estimation

	All manufacturing	Food and Beverages	Textiles and apparel	Paper and printing	Chemicals and Petroleum	Rubber and plastics	Metal products	Furniture
Infrastructure								
Market Access	-0.0008	-0.0012	0.0016	-0.0002	-0.0417	0.0007	0.0021	-0.0011
	[0.0003]**	[0.0007]*	[0.0012]	[0.0017]	[0.0193]**	[0.0019]	[0.0022]	[0.0007]
Electric grid	1.5043	2.3627	1.8944	12.7944	13.4268	15.5558	13.7753	0.6305
	[0.2846]***	[0.7139]***	[1.0224]*	[386.6443]	[1059.9987]	[1197.6892]	[345.9622]	[0.3322]*
Agglomeration								
Localization	-0.0007	0.0107	-0.0001	-0.0757	0.7126	0.1058	0.0043	0.0037
	[0.0002]***	[0.0025]***	[0.0006]	[0.0235]***	[0.3422]**	[0.0178]***	[0.0024]*	[0.0009]***
Diversity index	2.5932	2.9509	2.404	3.0168	0.4531	4.7781	3.4196	1.2262
	[0.2546]***	[0.4087]***	[0.9135]***	[1.1395]***	[3.996]	[1.1240]***	[0.8113]***	[0.4811]**
Human Capital								
Education	14.0667	5.5984	14.1216	26.4592	-14.6166	4.9415	13.0636	3.9464
	[0.4061]***	[1.3373]***	[1.6408]***	[2.8674]***	[14.0647]	[2.1058]**	[2.6103]***	[1.9758]**
Natural Geography								
Roughness	-0.0005	-0.0004	-0.0012	0.0002	0.0066	-0.0008	-0.0015	0.0002
	[0.0002]***	[0.0003]	[0.0006]*	[0.0005]	[0.0026]**	[0.0009]	[0.0006]**	[0.0002]
Observations	86562	28998	6642	8964	2322	4482	12150	21654

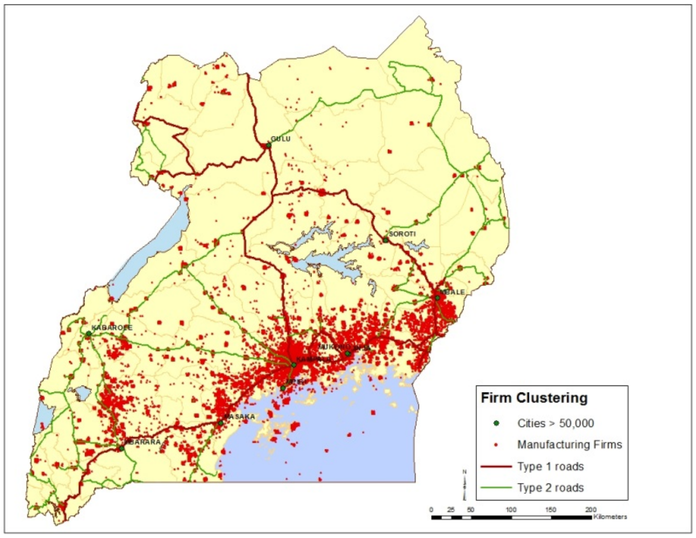
Standard errors in brackets

* significant at 10%; ** significant at 5%; *** significant at 1%

Table 2: Relative profits from conditional logit model (Kampala=100)

District Name	Relative Profits	District Name	Relative Profits
KAMPALA	100	KYENJOJO	50.8025
WAKISO	81.8655	BUSIA	50.7307
JINJA	74.7749	IGANGA	50.5119
MUKONO	66.8287	SIRONKO	50.2065
KALANGALA	66.321	RUKUNGIRI	49.7588
MUBENDE	61.6232	KAMULI	49.6589
HOIMA	61.2801	NAKASONGOLA	48.9634
KAPCHORWA	58.7859	RAKAI	48.3818
MBALE	57.7578	KUMI	48.1325
TORORO	57.595	MOROTO	47.916
LIRA	57.0166	KATAKWI	47.0133
KAYUNGA	56.9183	MAYUGE	46.8639
KABAROLE	56.9039	SEMBABULE	46.1921
LUWERO	55.9077	ARUA	45.7752
MASAKA	55.6198	KAMWENGE	45.4802
MPIGI	54.3784	NTUNGAMO	44.9891
BUSHENYI	54.048	KIBOGA	44.7983
MASINDI	53.4652	NEBBI	43.9336
MBARARA	53.0022	KOTIDO	43.6971
PALLISA	52.7676	ADJUMANI	41.9145
BUGIRI	52.5774	KABERAMAIDO	37.9408
GULU	52.4109	BUNDIBUGYO	34.9738
KASESE	52.3478	KANUNGU	34.6861
KIBAALE	52.1568	YUMBE	33.7736
MOYO	52.1228	NYAKAPIRIPIRITI	22.5532
APAC	51.6949	KISORO	18.1448
SOROTI	51.4554	KITGUM	NA
KABALE	50.8606	PADER	NA

Figure 1: Spatial distribution of manufacturing firms and Access to markets in Uganda



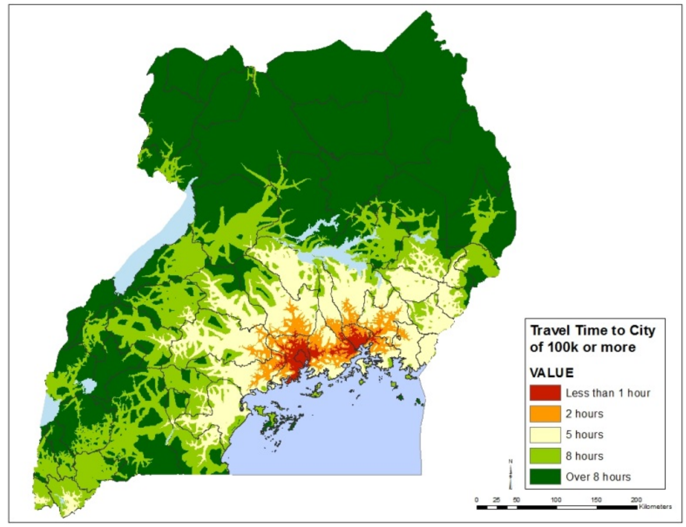


Figure 1: K density of overall manufacturing in Uganda

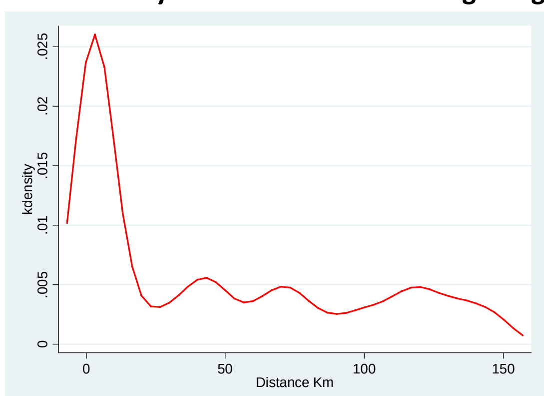


Figure 2: Comparing clustering across industries

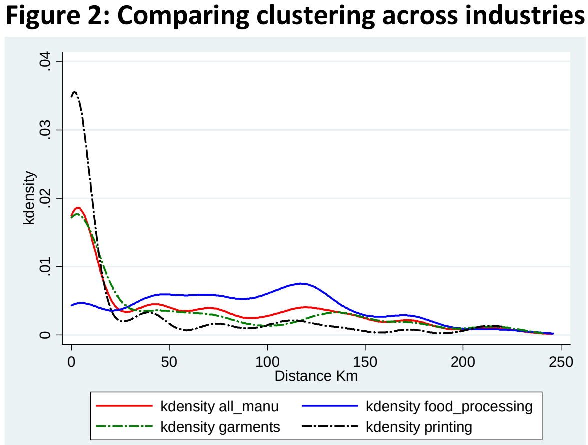
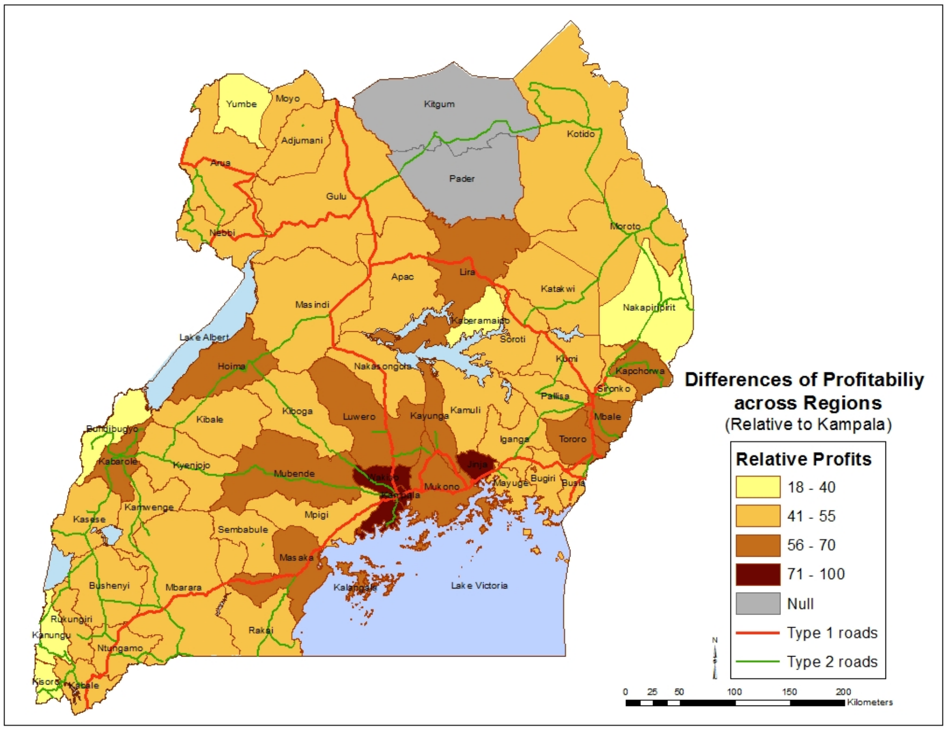
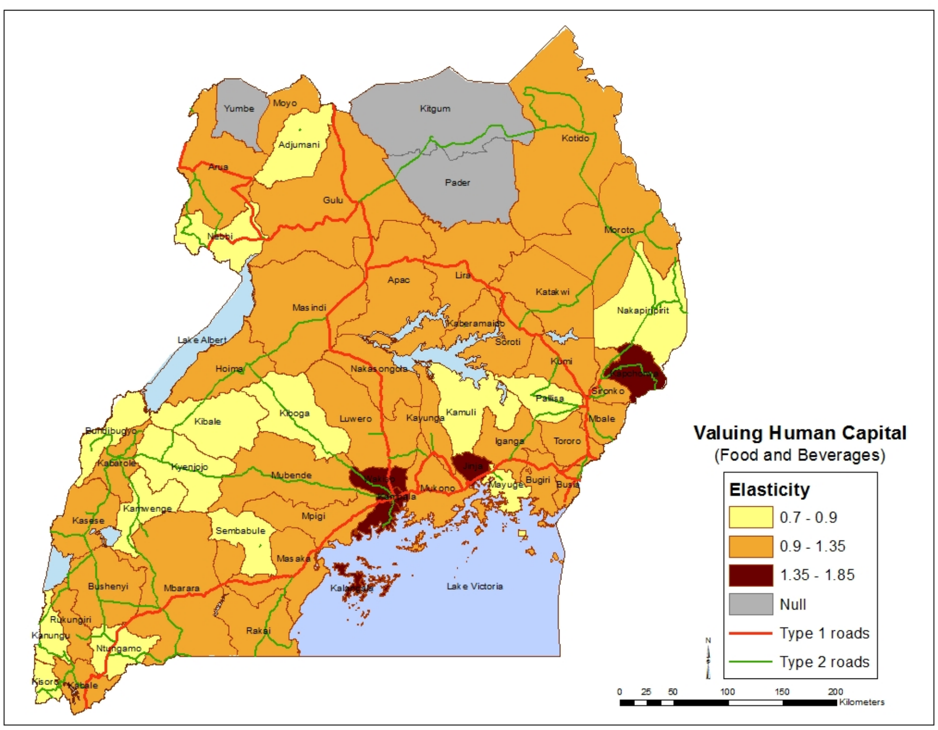


Figure 4: Relative profits for manufacturing firms across districts[a]



a. Profitability is compared relative to Kampala, which is normalized to 100

Figure 5: Importance of human capital for the food and beverage industry



ONLINE APPENDIX

DATA SOURCES AND VARIABLE CONSTRUCTION

Data sources

Travel time grid: Thomas (2007b); UNEP roads and national roads datasets, CIA World Data Bank II, Waterbodies data, and GRUMP/CIESIN global settlement points were used to create friction grid.

Roads dataset: UNEP roads are based on the Digital Chart of the World dataset. The World Bank added attributes (primary, secondary, and tertiary) from the 2004 Michelin map series.

Community points: Uganda National Household Survey, 2005/2006 (Uganda Bureau of Statistics, 2006).

Education: Uganda Population and Housing Census (Uganda Bureau of Statistics, 2002b)

Manufacturing business points: Uganda Business Registry (Uganda Bureau of Statistics, 2001 and 2002a).

Terrain roughness data: The CGIAR-CSI (Consortium for Spatial Information, http://srtm.csi.cgiar.org/), derived from the USGS/NASA SRTM data, Downloaded March 28, 2007.

Railroad dataset: USGS Global GIS Database: Africa CD, VMAP level 0, 2001.

Powerlines dataset: UGANDA-AERDP Final Report Annex 1; February 2004.

Crop Data : IFPRI Spatial Allocation Model; You and Wood (2003).

Variable construction

<u>Infrastructure for market access</u>: We computed travel times on the road network from each GIS pixel in the country to the nearest city of 100,000 or more people. In order to extract these data, we built a raster dataset at a 500 meter resolution where each pixel records the time in tens of minutes to travel from the specific pixel (year 2000 estimate from GRUMP alpha data). Travel time is estimated using a combination of several GIS layers that are merged into a friction grid which represents the time required to cross each pixel (each pixel represents $500m^2$). As the pixel friction value increases, the travel time to the nearest city of 100,000 increases as well.

The underlying road database is based on the Digital Chart of the World (DCW), which was expanded using attributes from Michelin 2004 regional maps to distinguish among primary, secondary, and tertiary roads. Thomas (2007b) assigns travel along primary roads to be 60km/hr, secondary roads 40km/hr, and tertiary roads 20km/hr. In addition, further refinements to these maps were made through a Uganda national roads dataset in order to allow for more precise analysis. This database was then layered with various GPS data in order to extract travel time and distance from major cities.

<u>Roughness</u>: Terrain roughness was derived from USGS/NASA SRTM data. In order to analyze roughness per district, these data were aggregated into 100 by 100 groups of cells (approximately 10 km by 10 km).

For each aggregation, a mean was computed in order to determine overall elevation variability within a district (Thomas, 2007a).

SENSITIVITY RESULTS

Table A1: Localization measured as own-industry concentration within 5 kilometers

	All manufacturing	Food and Beverages	Textiles and apparel	Paper and printing	Chemicals and Petroleum	Rubber and plastics	Metal products	Furniture
Infrastructure								
Market Access	-0.0020***	-0.0030***	-0.0000	-0.0041**	-0.0381*	0.0018	0.0009	-0.0022***
	[0.000]	[0.001]	[0.001]	[0.002]	[0.021]	[0.002]	[0.001]	[0.001]
Electric grid	1.4772***	2.3542***	1.7765*	13.7062	13.8956	16.0641	15.3712	0.6077*
	[0.283]	[0.714]	[1.016]	[575.259]	[2,269.575]	[1,391.947]	[809.204]	[0.331]
Agglomeration								
Localization	0.0079***	0.0049***	0.0158***	0.0134***	0.0040	-0.0057	0.0046*	0.0128***
	[0.001]	[0.002]	[0.004]	[0.004]	[0.006]	[0.004]	[0.003]	[0.002]
Diversity index	1.8894***	2.6067***	1.2466	1.1396	-2.8649	5.3869***	2.8306***	0.4856
	[0.264]	[0.412]	[0.938]	[1.187]	[4.892]	[1.207]	[0.869]	[0.494]
Human Capital								
Education	6.8555***	6.9386***	1.4658	6.3556*	7.1649	18.6964***	13.2594***	1.8226
	[0.849]	[1.371]	[3.505]	[3.562]	[8.957]	[3.544]	[2.487]	[1.718]
Natural Geography								
Roughness	-0.0002	-0.0003	-0.0007	0.0010*	0.0060**	-0.0019**	-0.0014**	0.0002
	[0.000]	[0.000]	[0.001]	[0.001]	[0.003]	[0.001]	[0.001]	[0.000]
Observations	86562	28998	6642	8964	2322	4482	12150	21654

Standard errors in brackets

* significant at 10%; ** significant at 5%; *** significant at 1%

Table A2: Localization measured using own-industry concentration within 60 kilometers

	All manufacturing	Food and Beverages	Textiles and apparel	Paper and printing	Chemicals and Petroleum	Rubber and plastics	Metal products	Furniture
Infrastructure								
Market Access	-0.0008*	-0.0018**	-0.0010	-0.0039*	-0.0454*	0.0060***	0.0023	-0.0011
	[0.000]	[0.001]	[0.002]	[0.002]	[0.025]	[0.002]	[0.002]	[0.001]
Electric grid	1.5078***	2.3774***	1.8362*	14.1837	12.7938	15.0388	14.5284	0.5151
	[0.285]	[0.715]	[1.023]	[808.525]	[1,324.930]	[846.672]	[501.902]	[0.334]
Agglomeration								
Localization	0.0000	0.0004	-0.0035**	-0.0037**	-0.0055	0.0051***	0.0006	-0.0001
	[0.000]	[0.001]	[0.001]	[0.002]	[0.005]	[0.001]	[0.001]	[0.001]
Diversity index	2.5890***	3.0412***	2.1129**	1.9971*	-6.2050	5.3388***	3.5361***	1.2709***
	[0.257]	[0.404]	[0.898]	[1.167]	[7.103]	[1.195]	[0.831]	[0.486]
Human Capital								
Education	13.3529***	10.4374***	18.3654***	23.0118***	16.4230**	9.3677***	16.8225***	11.9095***
	[0.491]	[0.775]	[2.212]	[2.449]	[6.671]	[1.709]	[1.432]	[0.986]
Natural Geography								
Roughness	-0.0005***	-0.0004	-0.0014**	0.0001	0.0058**	-0.0013	-0.0016**	-0.0001
	[0.000]	[0.000]	[0.001]	[0.001]	[0.003]	[0.001]	[0.001]	[0.000]
Observations	86562	28998	6642	8964	2322	4482	12150	21654

Standard errors in brackets

* significant at 10%; ** significant at 5%; *** significant at 1%

Table A3: Market access measured using distance to Kampala

	All manufacturing	Food and Beverages	Textiles and apparel	Paper and printing	Chemicals and Petroleum	Rubber and plastics	Metal products	Furniture
Infrastructure								
Market Access	-0.0005	-0.0014**	0.0015	0.0024	-0.0111	0.0026	0.0013	-0.0007
	[0.000]	[0.001]	[0.001]	[0.001]	[0.007]	[0.002]	[0.001]	[0.001]
Electric grid	1.5070***	2.2408***	1.9939*	13.5105	12.9522	15.9489	14.4602	0.5929*
	[0.290]	[0.720]	[1.037]	[411.447]	[1,809.970]	[1,204.987]	[492.671]	[0.344]
Agglomeration								
Localization	-0.0007***	0.0110***	-0.0001	-0.0802***	0.4625**	0.1060***	0.0042*	0.0036***
	[0.000]	[0.002]	[0.001]	[0.024]	[0.211]	[0.018]	[0.002]	[0.001]
Diversity index	2.6420***	2.8654***	2.4471***	3.5843***	2.8983	5.3286***	3.2677***	1.2217**
	[0.256]	[0.413]	[0.929]	[1.141]	[2.656]	[1.187]	[0.817]	[0.489]
Human Capital								
Education	14.2125***	5.1091***	14.2280***	28.8853***	1.1840	6.6681***	12.7939***	4.1749**
	[0.425]	[1.345]	[1.674]	[2.912]	[8.868]	[2.187]	[2.598]	[1.962]
Natural Geography								
Roughness	-0.0005***	-0.0003	-0.0012**	-0.0002	0.0035**	-0.0012	-0.0014**	0.0002
	[0.000]	[0.000]	[0.001]	[0.001]	[0.002]	[0.001]	[0.001]	[0.000]
Observations	86562	28998	6642	8964	2322	4482	12150	21654

Standard errors in brackets

* significant at 10%; ** significant at 5%; *** significant at 1%

Table A4: Including district output of coffee, cotton and maize as controls

	All manufacturing	Food and Beverages	Textiles and apparel	Paper and printing	Chemicals and Petroleum	Rubber and plastics	Metal products	Furniture
Infrastructure								
Market Access	-0.0025***	-0.0011	0.0003	-0.0019	-0.0340	0.0014	-0.0016	-0.0038***
	[0.000]	[0.001]	[0.002]	[0.002]	[0.021]	[0.003]	[0.002]	[0.001]
Electric grid	1.4578***	2.3294***	1.7155*	13.6191	14.5219	13.8397	15.2559	0.6341*
	[0.285]	[0.715]	[1.026]	[529.212]	[1,978.558]	[626.698]	[726.963]	[0.337]
Agglomeration								
Localization	-0.0009***	0.0111***	-0.0004	-0.0934***	1.0657*	0.0871***	-0.0004	0.0003
	[0.000]	[0.003]	[0.001]	[0.027]	[0.586]	[0.017]	[0.003]	[0.001]
Diversity index	2.7099***	2.8465***	2.5805**	3.5165***	-4.6242	4.4408***	3.7473***	1.2877**
	[0.271]	[0.427]	[1.073]	[1.145]	[8.209]	[1.401]	[0.858]	[0.513]
Human Capital								
Education	13.0736***	5.4191***	14.5239***	25.8973***	-16.1602	9.4507***	14.3492***	8.2810***
	[0.493]	[1.612]	[1.890]	[2.902]	[20.201]	[2.711]	[2.921]	[2.350]
Natural Geography								
Roughness	-0.0004**	-0.0004	-0.0012*	-0.0001	0.0068*	-0.0005	-0.0015**	0.0002
	-0.0004**	-0.0004	-0.0012*	-0.0001	0.0068*	-0.0005	-0.0015**	0.0002
Observations	86,562	28,998	6,642	8,964	2,322	4,482	12,150	21,654

Standard errors in brackets

* significant at 10%; ** significant at 5%; *** significant at 1%

Table A5: alternate specifications

	including interactions	mixed logit
Infrastructure		
Market Access	-0.0012*	-0.0022***
	[0.001]	[0.001]
Electric grid	2.3627***	2.3685***
	[0.714]	[0.714]
Agglomeration		
Localization	0.0107***	0.0007
	[0.003]	[0.001]
Diversity index	2.9509***	2.9868***
	[0.409]	[0.399]
Human Capital		
Education	5.5984***	10.5775***
	[1.337]	[0.658]
Natural Geography		
Roughness	-0.0004	-0.0005*
	[0.000]	[0.000]
Observations	85,212	85,212

Standard errors in brackets

* significant at 10%; ** significant at 5%; *** significant at 1%