Labor-Force Heterogeneity as a Source of Agglomeration Economies in an Empirical Analysis of County-Level Determinants of Food Plant Entry

Davis, David E. and Schluter, Gerald E.

South Dakota State University, Economic Research Service, USDA

December 2005

Journal of Agricultural and Resource Economics 30(3):480–501

Labor-Force Heterogeneity as a Source of Agglomeration Economies in an Empirical Analysis of County-Level Determinants of Food Plant Entry

David E. Davis and Gerald E. Schluter

Results of this study show that a heterogeneous labor force serves to attract new food manufacturing investment. We conduct analysis for SIC 20, Food and Kindred Product Manufacturing, and disaggregate analysis on all nine three-digit SIC food industries. Heterogeneity variables are a significant factor in nearly all specifications. We also examine which factors create the greatest increases in the expected number of new establishments. Areas with a high degree of labor heterogeneity are found to have large advantages. Labor heterogeneity is among the most important factors attracting food manufacturing to urban areas over rural areas.

Key words: agglomeration externalities, business location determinants, food manufacturing, labor heterogeneity, rural development

Introduction

Attracting new manufacturing plants is a key economic development strategy for many localities, and federal policies also support manufacturing as a development strategy for rural communities. For example, the 2002 Farm Bill establishes significant grants to facilitate the growth of manufacturing enterprises in rural areas (see, e.g., Title VI, *Rural Development* and Title IX, *Energy*). Developing a well-trained workforce is another strategy localities often employ to foster economic development *or* attract manufacturing plants. Worker training programs are far-ranging, and found in some form in every state. Vocational and training costs totaled $1.2 billion for state and local governments in 2001 (U.S. Department of Commerce, Bureau of the Census, 2004, table 524). The efficacy of funding training programs appears to be founded on a substantial literature that finds a more educated workforce attracts economic development. In particular, studies examining determinants of manufacturing plant growth and plant entry frequently recognize the importance of labor quality, and control for labor quality with a proxy variable—such as the percentage of an area's labor force with a high school degree.

A related literature suggests that the quality of a labor force, and hence its attractiveness to potential new investment, may be multi-dimensional. Firms may not only need

---

David E. Davis is assistant professor, Department of Economics, South Dakota State University, and former economist with the Economic Research Service, USDA. Gerald E. Schluter is an economist with the Economic Research Service, U.S. Department of Agriculture. The authors are grateful to Maureen Kilkenny, who provided helpful comments on earlier versions of this manuscript. The views in this article do not necessarily represent those of the Economic Research Service or the U.S. Department of Agriculture.

Review coordinated by T. F. Glover.

workers capable of learning new skills, but also may need a workforce with many different types of skills. A firm choosing to locate in an area with a deep labor pool represented by workers with various skills and training in a variety of occupations can expect lower worker training costs. Furthermore, if faced with worker turnover, firms may find the cost of replacing lost workers lower in areas with deep labor pools. A diverse labor market, therefore, can be a source of external scale economies for firms located in the market. External scale economies are thought to occur when groups of firms cluster together, reducing transportation and other costs and allowing for knowledge and training spillovers. Urban areas are favored by these external economies, a form of agglomeration economies in the tradition of Jacobs (1969) in which a diversity of knowledge or skills in a market reduces costs for firms.

While literature suggests that areas with heterogeneous labor pools may have an advantage in attracting manufacturing investment, few manufacturing plant location studies explicitly examine the issue. In this analysis, we extend the empirical research on the determinants of manufacturing plant location by testing for the effect of a heterogeneous local population and labor force. This study incorporates data from nearly all counties in the continental United States, and conducts analysis on SIC 20, Food and Kindred Product Manufacturing, as well as disaggregate analysis on all nine three-digit SIC food product categories (SIC 201–209) within SIC 20.[1] We find strong evidence that areas with heterogeneous labor markets attract more plants than areas with homogeneous labor markets. Because rural areas frequently have a more homogeneous labor force than other areas, this result has implications for rural development policy. Rural areas seeking to attract manufacturing plants will benefit by enacting policies that foster a labor force with a variety of education levels and occupations.

Plant Location and Labor-Force Heterogeneity

The relationship between labor-force heterogeneity and location is examined by Duranton and Puga (2001), who formalize the product cycle hypothesis first articulated by Jacobs (1969) and empirically examined by Glaeser et al. (1992) in which new products are developed in diversified cities. New product innovators borrow processes from other industries, thus requiring a labor force with diverse skills to accommodate the range of skills needed in the varying production activities. On finding their ideal process, firms switch to mass production and relocate to specialized cities where labor needs are more predictable and production costs are lower. Duranton and Puga find strong evidence of this pattern of economic activity in establishment relocations across French employment areas during 1993–1996. Implied in this analysis is that young industries most benefit from this external economy.

Kim (1990) also suggests a relationship between labor force diversity and plant location, and models agglomeration economies arising from heterogeneous labor markets.

[1] Food manufacturing is examined for two reasons. First, we suspect food manufacturing to be among the industries least likely to benefit from a heterogeneous labor force. Thus, a finding supporting our hypothesis should be met with a large degree of confidence. Second, we are interested in examining factors thought to benefit rural development. Because food manufacturing plants process raw farm products likely to be found in rural areas, they are thought more likely to locate in rural areas than other manufacturing industries (Henderson and McNamara, 2000).

As plants become more specialized in production, they are more likely to find workers with skills more closely matching their specialized needs in larger labor markets. Productivity increases and transactions costs decrease as the size of the labor market increases. One implication reported by Kim is that firms will be more likely to locate in areas with large labor markets "if the technology requires more specific labor, if its productivity is low, if the cost of training workers is high, or if the minimum efficient scale is large" (p. 176). Areas should be able to attract a larger number of firms, over a broader spectrum of industries, with a more diverse labor force.

In a study of new foreign-owned manufacturing plant locations, Coughlin and Segev (2000) included measures of racial diversity in a model of foreign-owned plant location and found a significant and positive effect. This result puzzled the authors because it had not been previously observed. As a potential explanation, we suggest the percentage of county population that is Black may be capturing the effect of labor-force heterogeneity.

Conceptual Framework

We follow previous research and use location theory to motivate our empirical model of plant locations (see, e.g., Bartik, 1985; Henderson, Kuncoro, and Turner, 1995; Henderson and McNamara, 2000; Coughlin and Segev, 2000; Woodward, 1992; and List, 2001). According to location theory, firms choose sites that maximize expected profits. This entails either minimizing delivered input costs, production costs, and costs related to marketing and delivering output, or optimizing output prices, or both. The decisions about which products to produce, for which markets, and at what scale, are assumed to be made prior to the decision of where to open the plant. Plants large enough to demand inputs from multiple states and to supply products to multiple markets will minimize costs by locating in an area central to their input or output markets. Small plants locate near their input or output market. Within that general area, firms choose a specific plant location.

Results from Guimarães, Figueiredo, and Woodward (GFW, 2003, 2004) suggest an appropriate estimation method. GFW (2003) contend that conditional logit models (CLMs) based on random utility (profit) maximization have been a fruitful method for modeling firm locations (for examples, see Bartik, 1985; Woodward, 1992). However, CLM models suffer from a limitation because of the difficulty estimating them with large choice sets and from an underlying independence of irrelevant alternatives (IIA) assumption. GFW show the equivalence between parameter coefficients estimated from the log likelihood of a fully specified CLM model and parameters estimated from the log likelihood of a Poisson regression. Specifically, parameter estimates from a Poisson regression under appropriate circumstances can be interpreted compatible with the random utility maximization framework. Extending their earlier result, GFW (2004) show that appropriately specifying a Poisson model can also control for the potential IIA violation common to conditional logit models.

Following GFW (2004), suppose there is an economy with k sectors, i investors, and j potential locations. The profit for investor i from selecting location j is assumed to be:

$$(1) \qquad \pi_{ijk} = \gamma'\mathbf{x}_k + \theta'\mathbf{y}_j + \beta'\mathbf{z}_{jk} + \varepsilon_{ijk},$$

where γ, θ, and β are vectors of unknown parameters; $\mathbf{x}_k$ is a vector of sector variables; $\mathbf{y}_j$ is a vector of location-specific variables; and $\mathbf{z}_{jk}$ is a vector of variables that vary with the region and sector. If $\varepsilon_{i,j}$ is an independently and identically distributed (i.i.d.) random error with Extreme Value Type 1 distribution, then the probability of selecting location j, conditional on choice of sector k, can be shown to equal:

$$p_{j/k} = \frac{\exp(\theta'\mathbf{y}_j + \beta'\mathbf{z}_{jk})}{\sum_{j=1}^{J} \exp(\theta'\mathbf{y}_j + \beta'\mathbf{z}_{jk})}. \tag{2}$$

This is the familiar conditional logit model. As demonstrated by GFW (2003), when n_{jk} denotes the number of investments in region j and sector k, then the parameters in equation (2) can be estimated from a Poisson model if n_{jk} follows a Poisson distribution with

$$E(n_{jk}) = \lambda_{jk} = \exp(\alpha_k + \theta'\mathbf{y}_j + \beta'\mathbf{z}_{jk}), \tag{3}$$

where α_k is a sector dummy variable. GFW (2003) establish a theoretical foundation based on a random utility framework for firm location studies using a Poisson regression. Yet to be resolved, however, is the underlying IIA assumption inherent in conditional logit models. GFW (2004) address this issue by adding an effect for each alternative to control for unobserved variables that affect firm location decisions, which they contend can cause a violation of the IIA assumption. Adding a term to the profit function in equation (1), we obtain:

$$\pi_{ijk} = \gamma'\mathbf{x}_k + \theta'\mathbf{y}_j + \beta'\mathbf{z}_{jk} + \gamma_j + \varepsilon_{ijk}. \tag{4}$$

If γ_j is a random variable, then, conditional on γ_j, the probability of choosing location j can be shown to equal:

$$p_{j/k\gamma} = \frac{\exp(\theta'\mathbf{y}_j + \beta'\mathbf{z}_{jk} + \gamma_j)}{\sum_{j=1}^{J} \exp(\theta'\mathbf{y}_j + \beta'\mathbf{z}_{jk} + \gamma_j)}. \tag{5}$$

Equation (5) is a variant of the mixed logit model, or a CLM with random effects. Given the association between the CLM and the Poisson regression, equation (5) can be estimated with a Poisson regression with random effects (GFW, 2004):

$$E(n_{jk}) = \lambda_{jk} = \exp(\alpha_k + \theta'\mathbf{y}_j + \beta'\mathbf{z}_{jk} + \gamma_j). \tag{6}$$

Now consider a single cross-section for a single sector, so there are no sector effects ($\mathbf{z}_{jk}$). Equation (6) becomes:

$$E(n_j) = \lambda_j = \exp(\alpha_0 + \theta'\mathbf{y}_j + \gamma_j). \tag{7}$$

If we assume $\exp(\gamma_j)$ is an i.i.d. gamma-distributed random variable with gamma parameters δ, δ^{-1}, so that $E(\exp(\gamma_j)) = 1$, and $Var(\exp(\gamma_j)) = \delta$, then equation (7) has a mixed Gamma/Poisson distribution which generates a negative binomial model (Cameron and Trivedi, 1998).

Our analysis is of a cross-section of plant locations for a single sector, Food Manufacturing (SIC 20), with choice-specific dependent variables. Because our dependent variable is a discrete nonnegative count of plant locations, we use a negative binomial regression as is suggested by equation (7), which (as shown by GFW, 2003) has the benefit of being compatible with a random utility (profit) maximization framework.[2]

Data and Variables

Descriptions of variables and their sources are provided in table 1. The dependent variable (*New Plants*) is a cumulative count of the number of new food manufacturing plants (SIC 20) locating in a county between 1991 and 1997. For each county, we count the number of establishments in county j in year t ($E_{j,t}$).[3] For county j in year t, if the number of establishments is greater than in year $t-1$, the number of new plants equals the difference in establishments; otherwise, the number of new plants is zero. This calculation is repeated for each year between 1991 and 1997, and then the number of new plants for all years is summed to obtain the total number of new plants for each county:[4]

$$(8) \qquad N_j = \sum_t N_{j,t},$$

$$\text{where } N_{j,t} = \begin{cases} E_{j,t} - E_{j,t-1} & \text{if } E_{j,t} > E_{j,t-1}, \\ 0 & \text{otherwise.} \end{cases}$$

This produces a cross-section of 3,111 observations, one for each county in the continental United States. We use previous research on plant locations to identify appropriate independent variables. Also following previous research, the analysis examines how exogenous variables prior to 1991 affect location decisions for 1991–1997 (e.g., Coughlin and Segev, 2000; Woodward, 1992; List and Kunce, 2000; List, 2001). The independent variables are 1990 values where possible. Where 1990 values are not available, we used the most recent data available prior to 1990.

Labor Market Characteristics

This study examines the effect of a heterogeneous labor force on plant entry. As in Coughlin and Segev (2000), we include the percentage of a county's population that is Black (*Black %*) as a potential indicator of a county's population and labor-force diversity.

[2] A potential shortcoming of our study is the inability to correct for possible spatial correlation in the errors of our regressions. When not controlled, spatial correlation renders coefficient estimates biased and inconsistent. Techniques are not currently available to estimate negative binomial regressions while simultaneously controlling for spatial correlation.

[3] Annual counts of establishments are from the County Business Patterns.

[4] This variable potentially undercounts the number of new plants in a county, because it is not possible to identify exiting establishments, which masks the entry of some new plants. Difficulty measuring gross entry is a common problem in economic studies. Industrial organization economists face a similar problem when examining gross entry and market structure. Using net entry as a proxy is a frequently used solution (see, e.g., Orr, 1974; Duetsch, 1984; Chappell, Kimenyi, and Mayer, 1990). Based on findings reported by Dunne, Roberts, and Samuelson (1988), this method seems justified. Specifically, they found that when entry rates are high, exit rates are low—suggesting gross entry and net entry are correlated. We contend, because our measure of entry is narrowly defined over time (we calculate annual changes in establishment counts, then sum these annual counts) and over space (we calculate each change at the county level), it is less likely to encounter instances in which entry and exit occur simultaneously. Given this, "true" new plant entry into a county should be captured quite accurately.

Table 1. Variable Names, Description, and Sources

Variable	Description	Source
New Plants	Number of new plants entering a county between 1991 and 1997 (dependent variable)	County Business Patterns
High School %	Percent of county population age 25 and over with at least a high school degree	U.S. Bureau of the Census
Union %	Percent of state manufacturing employment that is unionized	*Statistical Abstract of the U.S.*
Population	County population	U.S. Bureau of the Census
Pop Squared	County population squared	
New England	New England states: CT, ME, MA, NH, RI, VT = 1; other states = 0	
Mideast	Mideast states: DE, MD, NJ, NY, PA = 1; other states = 0	
Great Lakes	Great Lakes states: IL, IN, MI, OH, WI = 1; other states = 0	
Plains	Plains states: IA, KS, MN, MO, NE, ND, SD = 1; other states = 0	
Southwest	Southwest states: AZ, NM, OK, TX = 1; other states = 0	
Rocky Mountain	Rocky Mountain states: CO, ID, MT, UT, WY = 1; other states = 0	
Far West	Far West states: CA, NV, OR, WA = 1; other states = 0	
Southeast	Southeast states: AL, AR, FL, GA, KY, LA, MS, NC, SC, TN, VA, WV = 1; other states = 0	
Manuf Empl %	1990 county manufacturing employment/ labor force	U.S. Bureau of the Census
Black %	1990 percent of county population that is Black	U.S. Bureau of the Census
Hispanic %	1990 percent of county population of Hispanic descent	U.S. Bureau of the Census
Gini	GINI ratio	Authors' calculations from *1990 Census*
Education HHI	County education Herfindahl index	Authors' calculations from *1990 Census*
Occupation HHI	County occupation Herfindahl index	Authors' calculations from *1990 Census*
Per Capita Prop Tax	County 1987 per capita property tax	USA Counties
Wages/Value Added	1987 production worker hourly wage/value added per hour	U.S. Bureau of the Census
Population Density	1990 county population per square mile	U.S. Bureau of the Census
Per Capita Income	1990 total county personal income/county population	Bureau of Economic Analysis
Unemployed	1990 county unemployment rate	U.S. Bureau of the Census
Highway	Counties with an interstate highway = 1; 0 otherwise	ArcView 4.3
Right-to-Work	County located in right-to-work state = 1; 0 otherwise	*Statistical Abstract of the U.S.*
Manuf Estab	1990 county manufacturing establishments	Authors' calculations from County Business Patterns
Value of Crops	1987 value of crops produced in a county	U.S. Bureau of the Census
Value of Livestock	1987 value of livestock produced in a county	U.S. Bureau of the Census

We do not necessarily believe food manufacturing plants target locations with large (or small) Black populations. Instead, our hypothesis is that areas with heterogeneous labor forces are also likely to be racially diverse. So, a regression of new plants on measures of racial diversity is likely to identify a relationship, but the relationship arises from a correlation between racial diversity and labor-force heterogeneity.

A recent change in the ethnic makeup of food manufacturing labor suggests inclusion of another race variable. Because Hispanic workers represent an increasingly large proportion of the workforce at food plants, we include the percentage of a county's population of Hispanic descent (*Hispanic%*) as an additional variable.

Testing our hypothesis requires variables that measure labor-force heterogeneity. We are not aware of any previous studies that measure labor-force heterogeneity, and therefore have little guidance for appropriate measures. Instead, we offer some plausible variables for heterogeneity including measures of educational heterogeneity, occupational heterogeneity, and income inequality. Educational heterogeneity should be directly related to skill heterogeneity since higher educated workers are also more skilled workers. As a proxy for educational heterogeneity, a Herfindahl-Hirschman index (HHI) is calculated from the population shares of persons 25 years of age and over who have completed various levels of education (*Education HHI*). The *1990 Census* reports six different levels of educational attainment for persons 25 years of age and older, which range from having less than a 9th-grade education to having a graduate or professional degree. The educational HHI for county i is

$$HHIEDU_i = \left(\sum_j s_{i,j}^2\right) * 10,$$

where $s_{i,j}$ is the share of persons over age 25 in county i with education level j. Herfindahl-Hirschman indexes are frequently used to measure diversity or heterogeneity (see, e.g., Henderson, Kuncoro, and Turner, 1995). An increase in *HHIEDU* indicates a decrease in educational heterogeneity. If a county's population over age 25 were evenly distributed among the six education levels, then $HHIEDU_i$ equals 1.67; if it is concentrated in a single level, then $HHIEDU_i$ equals 10 (HHI is multiplied by 10 to improve convergence properties).

A similar measure is included to proxy for occupational heterogeneity (*Occupation HHI*). The *1990 Census* reports the number of persons 16 years of age and older employed in 13 different occupation categories. We calculate an HHI from the shares of persons in each of these categories, which range from "executive, administrative, and managerial occupations" to "handlers, equipment cleaners, helpers, and laborers."

Income inequality is measured with a county-level Gini coefficient. A Gini coefficient measures the distance between the cumulative distribution of a population's income from a uniform or equal distribution. An increase in a Gini coefficient indicates an increase in income inequality and suggests heterogeneity in wage earners—i.e., inequality implies a mix of high- and low-wage earners.

Our hypothesis is that counties with more diverse populations in terms of education and occupation are relatively more attractive to new plants than less diverse counties. Because lower HHI values indicate a more evenly distributed or more diverse population, we expect negative coefficients on these variables. In contrast, a higher value for a Gini coefficient indicates a more unequal distribution of income, and thus suggests a population more diversely represented by high and low income levels, so a positive sign is expected for the Gini coefficient.

A number of controls for labor market conditions are included, as suggested by previous research. A firm's assessment of labor costs in a potential location should compare wage costs relative to the productivity of the workers earning those wages. In our model, we include a productivity-adjusted measure of county-level hourly wages earned by production workers (Coughlin and Segev, 2000). Average hourly production worker wages are divided by the hourly value added (*Wages/Value Added*). If either county-level hourly wages or county value added were not available, productivity-adjusted wages were calculated at the state level, and this value was substituted for the missing county data. Higher productivity-adjusted wages are expected to be less attractive to prospective firms. Coughlin and Segev (2000) used a similar measure and found a negative relationship between wages and new foreign-owned manufacturing plants. Goetz (2000) found a negative relationship between food plant growth and wages, and Henderson and McNamara (2000) reported a negative relationship between food plant investments and wages.

This analysis controls for labor-force quality with a commonly used measure of educational attainment, the percentage of county population over 25 years of age with a high school diploma (*High School %*), which serves as a proxy for labor quality (see, e.g., Henderson and McNamara, 2000; Coughlin and Segev, 2000). Woodward (1992) used the median year of school completed as a measure of education attainment. Each of these studies found a positive effect.

We control for a number of other labor market conditions, including county-level unionization (*Union %*), and whether a county is located in a right-to-work state (*Right-to-Work*). Union workers receive higher wages, and frequently allow management less flexibility. Thus, a high percentage of union workers may make a county a less attractive location; Bartik (1985) and Woodward (1992) found evidence consistent with this hypothesis. Coughlin and Segev (2000) included a right-to-work dummy variable, but did not find a significant effect. We suggest that right-to-work legislation likely affects union security and strength (Hogler, 1995, p. 207). An area's attractiveness to potential investors may be related to union strength, and the percentage of the workforce unionized may capture only one aspect of union strength; right-to-work laws are likely to capture another.

We control for local labor-market conditions with the county unemployment rate (*Unemployed*). If a higher rate indicates a higher level of labor availability, and has a dampening effect on wages, then a positive influence is anticipated. Goetz (2000) found a positive relationship between unemployment rate and food plant growth for some industries. If, however, unemployment indicates a poor economic environment, and a less favorable quality of life, then a negative influence is expected, as was found by Henderson and McNamara (2000) and Woodward (1992).

As in previous studies, we include measures of agglomeration which are now commonly grouped into two classes. Localization economies are externalities that arise from a group of firms producing a similar product in close proximity, while urbanization economies are defined as externalities associated with a high level of overall economic activity located in a particular area. To control for localization economies, we include the percentage of the labor force employed in manufacturing (*Manuf Empl %*) and the number of manufacturing establishments located in a county (*Manuf Estab*) in 1990. We control for urbanization economies with county population (*Population*) and population density (*Population Density*). Goetz (2000) found population to positively affect food plant growth for some food manufacturing industries, but to negatively affect others. We suspect population may affect location decisions in a nonlinear manner, and therefore include a quadratic on population (*Pop Squared*); urbanization economies may diminish as population increases, since congestion decreases urbanization benefits. Population density may also proxy for land costs. Population is hypothesized to positively influence location decisions, while the effect of population density may be positive or negative.

Other Variables

Government policies may affect firm location decisions; higher tax levels may indicate higher business costs and thus deter entry. However, higher tax levels may also indicate higher levels of public goods and services, such as education, training, and infrastructure. Following Coughlin and Segev (2000) and Bartik (1985), we include per capita property tax (*Per Capita Prop Tax*). Coughlin and Segev did not find a significant effect, whereas Bartik reports a significant and negative relationship. Thus, we also include a highway dummy variable (*Highway*) as a proxy for transportation infrastructure availability, which takes a value of one when an interstate highway is present in a county, and zero otherwise.

Controls for both input supply and output demand are included. Because food manufacturing plant costs are sometimes dominated by raw agricultural material input costs, we include controls for input availability (Henderson and McNamara, 2000; MacDonald et al., 2000). We incorporate measures of the total value of crops produced in a county in 1987 (*Value of Crops*), and the total value of livestock produced in a county in 1987 (*Value of Livestock*). Henderson and McNamara found a positive association between food plant investments and crop and livestock production, and we also expect a positive coefficient for these variables. Moreover, we control for access to output markets with per capita county income in 1990 (*Per Capita Income*).

Some suggest firms initially determine a region within which to locate, and then choose a specific site within that region (Schmenner, Huber, and Cook, 1987). If location characteristics affecting the upper-level decision are different than those at the lower level, then it is important to include those characteristics in the model. As in previous studies, we include Bureau of Economic Analysis (BEA) regional dummy variables to represent these unobservable factors (refer to table 1 for a state-by-state affiliation).

Because support for manufacturing is frequently touted as a rural development strategy, controls for metropolitan location are incorporated. We create dummy variables based on Beale code classifications (table 2). Beale codes create an urban-to-rural continuum variable for U.S. counties. Their values range from 0 (the most urban) to 9 (the most rural). For counties with Beale codes of 3, 4, or 5, we define a suburban dummy variable (*Suburban*) that takes a value of one, and zero otherwise. For counties with Beale codes of 6, 7, 8, or 9, we define a rural dummy variable (*Rural*) that takes a value of one, and zero otherwise.

Table 3 presents the mean values and standard deviations for all variables. Urban counties attracted the highest number of new food plants. Although food manufacturers are thought to be more likely than other manufacturers to locate in rural areas, these data suggest that most new food manufacturing plants locate in urban counties. Rural counties have nearly as large a percentage of their workforce employed in manufacturing (17%) as urban counties (18%), but urban counties have many more manufacturing

Table 2. Beale Code Classifications

Code No.	Description
0	Central counties of metro areas of 1 million population or more
1	Fringe counties of metro areas of 1 million population or more
2	Counties in metro areas of 250,000 to 1 million population
3	Counties in metro areas of fewer than 250,000 population
4	Urban population of 20,000 or more, adjacent to a metro area
5	Urban population of 20,000 or more, not adjacent to a metro area
6	Urban population of 2,500 to 19,999, adjacent to a metro area
7	Urban population of 2,500 to 19,999, not adjacent to a metro area
8	Completely rural or less than 2,500 urban population, adjacent to a metro area
9	Completely rural or less than 2,500 urban population, not adjacent to a metro area

Source: Codes developed by USDA/Economic Research Service.

Table 3. Variable Means and Standard Deviations

	All Counties				
Variable	Mean	Standard Deviation	Urban Mean	Suburban Mean	Rural Mean
New Plants	2.270	3.98	5.510	3.030	1.110
Urban	0.202	0.40	1.000	0.000	0.000
Suburban	0.145	0.35	0.000	1.000	0.000
Rural	0.653	0.48	0.000	0.000	1.000
High School %	69.538	10.34	75.041	73.693	66.912
Union %	18.464	11.68	21.886	19.263	17.226
Population (10,000s)	7.930	26.50	28.040	7.850	1.710
Pop Squared	762.531	15,173.52	3,696.284	80.545	4.681
New England	0.022	0.15	0.035	0.040	0.013
Mideast	0.057	0.23	0.151	0.069	0.025
Great Lakes	0.141	0.35	0.180	0.169	0.122
Plains	0.199	0.40	0.070	0.142	0.251
Southwest	0.122	0.33	0.091	0.129	0.130
Rocky Mountain	0.069	0.25	0.021	0.049	0.089
Far West	0.048	0.21	0.072	0.080	0.034
Southeast	0.343	0.48	0.382	0.322	0.335
Manuf Empl %	17.321	9.88	18.036	17.664	17.024
Black %	8.626	14.37	10.284	8.514	8.137
Hispanic %	4.481	11.10	4.880	4.912	4.262
Gini	41.737	3.49	39.856	41.687	42.331
Education HHI	2.204	0.27	2.062	2.080	2.276
Occupation HHI	1.140	0.20	1.157	1.108	1.142
Per Capita Prop Tax	425.582	354.66	409.504	357.571	445.639
Wages/Value Added	0.264	0.10	0.233	0.246	0.277
Population Density	218.521	1,428.09	851.403	162.369	34.868
Unemployed	6.646	3.07	5.848	6.684	6.885
Per Capita Income	17.505	57.25	27.276	17.561	14.465
Highway	0.439	0.50	0.820	0.607	0.284
Right-to-Work	0.543	0.50	0.450	0.496	0.583
Manuf Estab	11.388	40.88	38.809	12.051	2.745
Value of Crops	1.875	4.68	3.005	2.529	1.379
Value of Livestock	2.470	4.23	2.314	3.397	2.313
N	3,109		629	450	2,030

plants on average. This result may explain the emphasis of developing manufacturing as a rural development strategy. While relatively few manufacturing plants locate in rural areas, with a small rural labor force, relatively few plants are needed for manufacturing employment to constitute as large a proportion of total employment as in urban areas.

Results

Table 4 presents the results from estimating three negative binomial regression models for SIC code 20, Food and Kindred Product Manufacturing, using data from a cross-section of 3,109 U.S. counties.[5] We begin by attempting to replicate the positive coefficient for proportion of county population that is Black (*Black %*) as was found in Coughlin and Segev (2000). A likelihood-ratio test shows that the model is significant at the 1% level. To measure the fit of the model, we calculate a measure suggested by Cameron and Windmeijer (1996), i.e., $R^2_{DEV,NB2}$, which is calculated as:

$$R^2_{DEV,NB2} = 1 - \frac{\sum_{i=1}^{n}\left\{y_i \log\left(\frac{y_i}{\hat{\lambda}_i}\right) - (y_i + \hat{\alpha}^{-1})\log\left(\frac{(y_i + \hat{\alpha}^{-1})}{\hat{\lambda}_i + \hat{\alpha}^{-1}}\right)\right\}}{\sum_{i=1}^{n}\left\{y_i \log\left(\frac{y_i}{\bar{y}_i}\right) - (y_i + \hat{\alpha}^{-1})\log\left(\frac{(y_i + \hat{\alpha}^{-1})}{\bar{y}_i + \hat{\alpha}^{-1}}\right)\right\}},$$

and is reported in table 4.

The Effect of Heterogeneous Labor Markets

The first specification in table 4 presents results without heterogeneity variables, whereas specification 2 includes the heterogeneity variables.[6] In specification 1, the coefficient for *Black %* is positive and significant, consistent with the findings of Coughlin and Segev (2000). In specification 2 (with heterogeneity variables), the sign on the *Black %* coefficient changes from positive to negative, and it is no longer significant. Meanwhile, the coefficient estimates for the heterogeneity variables have the expected signs, and are statistically significant. These results support our hypothesis that labor-force heterogeneity is a mechanism which attracts food manufacturing plants and not high proportions of Black workers. In previous research on determinants of Japanese plant locations (Woodward, 1992), Black population was found to have a negative impact on the likelihood of a plant locating in a county. In contrast, Smith and Florida (1994) report a positive association between Japanese auto-related manufacturing locations and

---

[5] Insufficient data require we omit two observations from our data: FIPS code 11001 (District of Columbia) and FIPS code 30113 (North Yellowstone, MT).

[6] Attention must be paid in location studies to potential endogeneity bias. In our study, several variables are potentially simultaneously determined with plant entry. Our solution is to follow the convention in plant location studies and use values of right-hand-side variables from a period prior to entry (i.e., right-hand-side variables are from t, whereas new plants enter from $t + 1$ to $t + 7$). Hence, these variables are predetermined and exogenous unless entry is anticipated, or plant construction times are long. We tested for bias from the latter source with a Hausman (1978) exogeneity test. The variables *Black %*, *Hispanic %*, *Occupation HHI*, *Education HHI*, *Gini*, *Value of Crops*, and *Value of Livestock* were treated as potentially endogenous. As instruments, we used 1982 values for *Value of Livestock* and *Value of Crops*, and 1980 values for the other variables. The χ^2 test statistic, with 32 degrees of freedom, is 0.45, and the null of no significant bias could not be rejected.

Table 4. Parameter Estimates from Negative Binomial Regression Models for SIC 20, Food and Kindred Product Manufacturing (3,109 U.S. counties)

	Specification 1		Specification 2		Specification 3	
Variable	Parameter	Std. Error	Parameter	Std. Error	Parameter	Std. Error
Constant	-1.240***	0.291	0.144	0.609	0.325	0.612
Black %	0.004**	0.002	-0.001	0.002	-0.001	0.002
Hispanic %	0.009***	0.002	0.008***	0.002	0.008***	0.002
Gini			0.023***	0.008	0.024***	0.008
Education HHI			-0.500***	0.087	-0.508***	0.088
Occupation HHI			-1.126***	0.163	-1.105***	0.164
Suburban	-0.018	0.053	-0.105*	0.054	-0.101*	0.054
Rural	-0.519***	0.056	-0.545***	0.057	-0.535***	0.057
High School %	0.030***	0.003	0.030***	0.003	0.030***	0.003
Union %	-0.002	0.003	-0.003	0.003	-0.007**	0.003
Population	0.008***	0.002	0.010***	0.002	0.010***	0.002
Pop Squared	-1.8E-05***	1.5E-07	-1.7E-05***	1.4E-07	-1.7E-05***	1.4E-07
Manuf Empl %	0.005**	0.003	0.007***	0.003	0.008***	0.003
Per Capita Prop Tax	-3.1E-04***	7.5E-05	-2.2E-04***	7.5E-05	-2.4E-04***	7.6E-05
Wages/Value Added	-0.569***	0.212	-0.581***	0.209	-0.553***	0.209
Population Density	8.7E-06	1.0E-05	1.0E-05	9.9E-06	9.7E-06	9.8E-06
Unemployed	-0.008	0.009	-0.024***	0.009	-0.023*	0.009
Per Capita Income	-7.3E-04	4.7E-04	-4.6E-04	4.7E-04	-4.9E-04	4.8E-04
Highway	0.145***	0.039	0.124***	0.039	0.126***	0.039
Right-to-Work (RTW)	-0.138**	0.057	-0.136**	0.056	-0.370***	0.100
RTW × Union %					0.019***	0.007
Manuf Estab	0.006***	0.001	0.004***	0.001	0.004***	0.001
Value of Crops	0.007**	0.003	0.004	0.003	0.004	0.003
Value of Livestock	0.020***	0.004	0.017***	0.004	0.016***	0.004
New England	0.447***	0.119	0.425***	0.117	0.405***	0.117
Mideast	0.173*	0.101	0.259**	0.101	0.347***	0.106
Great Lakes	-0.058	0.091	0.048	0.092	0.126	0.096
Plains	-0.117*	0.070	-0.053	0.070	-0.044	0.070
Southwest	-0.026	0.078	-0.132*	0.078	-0.153*	0.078
Rocky Mountain	0.002	0.096	0.052	0.096	0.071	0.096
Far West	0.324***	0.096	0.300***	0.094	0.326***	0.095
alpha (α)	0.273***	0.020	0.248***	0.019	0.247***	0.019
Log Likelihood	-5,178.76		-5,171.20		-5,167.24	
$R^2_{DEV,NB}$	0.507		0.526		0.527	

Note: Single, double, and triple asterisks (*) denote statistical significance at the 0.10, 0.05, and 0.01 levels, respectively.

concentrations of minority workers. The authors of these papers suggest that managers may have a preference for areas with or without large proportions of minority workers. However, we suggest that managers are acting in a profit-maximizing manner. By seeking areas with heterogeneous labor forces, training and turnover costs are reduced.[7]

Interestingly, the coefficient for *Hispanic %* is positive and significant in both specifications 1 and 2, retaining explanatory power even when heterogeneity variables are included. We offer the following potential explanations for consideration. First, food manufacturing plants may indeed target areas with large Hispanic populations. Food manufacturing jobs are often physically taxing, dirty, and dangerous. Historically, new immigrant populations have been willing to accept an economy's least desirable jobs (Hopkins, 2003). Hispanics are the newest wave of immigrants into the United States, and food plants may target areas with high Hispanic populations recognizing their willingness to take jobs hard to fill in other areas. Second, the positive association may represent a form of environmental racism. Food plants can be large, sprawling, unattractive, or otherwise unappealing enterprises. Firms may anticipate, or face, opposition from local populations when choosing where to locate. Minority or low economic status groups may have less power to oppose a plant locating in close proximity to them, or may feel economic pressure to accept them in exchange for the income and jobs arising from a new plant.[8]

Interpreting Heterogeneity and Labor Market Coefficients

We begin by examining labor market variables and noting a surprising result in specification 2. The initial coefficient estimates for unionization do not seem to have a significant impact for attracting new plants. However, the coefficient for the *Right-to-Work* dummy variable is negative and statistically significant. These are somewhat surprising results—a negative effect was expected from unionization and a positive effect from right-to-work. Consequently, we investigated them further. It is possible that unionization is important in plant location decisions only if unions are strongly organized. Right-to-work laws likely reduce union strength, as workers are not compelled to join unions in such jurisdictions. If so, the relationship between unionization and plant locations may not be linear, but instead unionization may have a different effect depending on whether a state has a right-to-work law. We tested this hypothesis by including an interaction variable (*RTW* × *Union %*), which is percentage unionized multiplied by the right-to-work dummy variable. Results are reported in specification 3 of table 4. When the interaction term is included, the unionization variable (*Union %*) is negative and significant, the *Right-to-Work* variable is negative and significant, but the interaction coefficient (*RTW* × *Union %*) is positive and significant. Interpreting the coefficients, we observe that increases in unionization imply increases in the expected

[7] We experimented with a "fractionalization index," suggested by Alesina, Baqir, and Easterly (1999), as a measure of racial diversity. The fractionalization index equals one minus the sum of the squared shares for each race. Substituting this variable instead of *Black %* and *Hispanic %* resulted in a significant coefficient when heterogeneity variables were not included. The coefficient was not significant when heterogeneity variables were included, further supporting our argument that heterogeneity is more important than racial diversity.

[8] We are indebted to an anonymous reviewer for suggesting this explanation. While we did not examine the role of environmental considerations, Adhikaril, Harsh, and Cheney (2003) did in their attempt to explain the regional shifts of U.S. pork production. Reporting their findings, they note, "Environmental compliance cost is considered one of the major factors of industry relocation; the analysis showed that the effect of such costs was minimal."

number of new plants when a county is in a right-to-work state, but decreases when a county is not in a right-to-work state. Prior research has found conflicting results with regard to unionization and plant locations. Bartik (1985) reports evidence that higher levels of unionization serve to dissuade firms from locating in an area. In contrast, Friedman, Gerlowski, and Silberman (1992) found a positive association between plant location and unionization. Our result may shed light on the apparent disagreement in earlier findings. We suggest that unions are seen as productivity-enhancing when weakly organized—as is likely the case in right-to-work states. On the other hand, when strongly organized—as is likely in states without right-to-work laws—unions represent a threat and are seen as an impediment to profit maximization for new plants.

We interpret the magnitude for continuous variables by referring to table 5, which reports the results of a one standard deviation change in continuous variables using coefficient estimates from specification 3 in table 4.[9,10] We observe that dispersion in the labor-force heterogeneity variables has a large effect.[11] The Gini coefficient has a standard deviation of 3.49 (table 3), which is about 8% of the variable's mean, and a one standard deviation change in income inequality suggests an 8.22% change in the expected number of new plants (table 5). The standard deviation for educational heterogeneity (*Education HHI*) is 0.27 in our data (table 3), which implies a 13.93% change in expected new plants (table 5). A similar calculation for occupational heterogeneity (*Occupation HHI*) implies a 22.5% change in the expected number of new plants from a one standard deviation change. In contrast, a standard deviation change in percentage of high school graduates (*High School %*) implies a 30.99% change in expected new plants. While educational attainment is an important attractor for new food plants, a heterogeneous labor force is also important. A county with a homogeneous labor force seems at a significant disadvantage.

Interpreting Other Coefficients

Other labor market variables are also among the most important determinants. As expected, counties with high wages relative to productivity are at a disadvantage.[12] A one standard deviation increase in this wage variable (*Wages/Value Added*) implies a 5.48% decrease in expected new plants. Unemployment rates appear to indicate that other labor market characteristics are more important than the level of labor availability. The negative coefficient on *Unemployed* suggests food plants avoid areas with high rates of unemployment. Woodward (1992) also found this result.

---

[9] An elasticity estimate for a continuous variable is equal to its coefficient estimate multiplied by its mean. We do not report these results here, and instead leave these for the interested reader to calculate from tables 3 and 4.

[10] Negative binomial regression coefficients for continuous variables represent a proportionate change in the conditional mean of the dependent variable from a one-unit change in an independent variable. The exponent of coefficients for dummy variables (e.g., $\exp(dj)$) implies the conditional mean is $\exp(dj)$ times larger when the dummy variable is one rather than zero (Cameron and Trivedi, 1998).

[11] The simple correlation coefficient between the education HHI and the high school percentage is -0.19, and between the occupation HHI and high school percentage is 0.14. Similarly, the correlation between the occupation HHI and education HHI is -0.07.

[12] We also tested whether a less restricted treatment of *Wages/Value Added* affected results. Rather than dividing wages by value added, we estimated an unrestricted model in which each variable (wages, value added) was included as a separate variable. In this specification, the coefficient for wages was negative, the value-added coefficient was positive, both were significant, and other values were largely unchanged. However, the restricted model provided the better fit based on a Bayesian information criterion.

Input supply variables seem to play a relatively small role in attracting food plants. Estimates in table 5 show that a one standard deviation change in *Value of Crops* results in only a 1.86% change in expected plants, while the *Value of Livestock* estimate suggests a 6.86% change. Also somewhat surprising is that *Per Capita Income* is not a statistically significant determinant.

In contrast, market size is an important determinant and a one standard deviation increase in *Population* increases expected new plants by 27.19%.[13] And localization economies also seem important, as a standard deviation increase in manufacturing employment (*Manuf Empl %*) equates to a 7.45% increase in expected new plants. Meanwhile, a standard deviation change in the number of manufacturing establishments (*Manuf Estab*) implies a 16.92% change in expected new plants.

Public-sector variables have the expected effects. Per capita property taxes (*Per Capita Prop Tax*) has a negative effect, and a one standard deviation change is estimated to imply an 8.4% decrease in the expected number of plants (table 5). Referring to table 4, counties with interstate highways attract 13.5% more plants than counties without them.

Urban versus Rural

Although some speculate that rural areas have an advantage attracting food manufacturing over other industries, our results suggest these areas attract fewer food plants than urban areas. Ceteris paribus, the coefficient for rural counties (table 4) implies the expected number of new plants is only 59% of the expected number in urban counties. Suburban counties also seem at a disadvantage, but to a lesser degree in that they attract about 10% fewer plants than urban counties.

As shown in table 3, on average urban counties attract more plants than rural counties. Using mean values from table 3 and coefficients from specification 3 in table 4, we can get a sense of the important factors providing advantage to urban areas. Coefficients are multiplied by the difference in the mean values for all continuous variables and results are summarized in table 6. Educational attainment is an important factor in generating differences in the number of plants choosing urban over rural locations. The percentage of the population over age 25 who are high school graduates (*High School %*) average about 8% less in rural counties than urban counties, which translates into a 24.35% reduction in the expected number of new plants in rural areas.

Some theorize that rural areas offer a workforce with a limited number of skills, and that these "one-note" labor forces may be an impediment to attracting manufacturing. Our results with regard to educational heterogeneity (*Education HHI*) seem to confirm this notion. A greater heterogeneity in educational attainment for urban areas implies a 10.85% advantage in expected new plants over rural areas. Rural areas have a small advantage in average occupational heterogeneity (*Occupation HHI*), which translates into a small reduction in the difference in average new plants between urban and rural areas. Rural areas have greater income inequality, but the difference results in a relatively small reduction in urban/rural new plant difference.

---

[13] Most location studies do not include *Pop Squared*, and we estimated the model without this variable to observe the effect on results. They were largely unchanged. However, in the specification without *Pop Squared*, the magnitude (in absolute value) of the *Suburban* and *Rural* coefficients increases by 0.09 and 0.10, respectively. Furthermore, the *Population Density* coefficient is larger (a one standard deviation change implies a 3.27% change in expected new plants) and statistically significant in the model without *Pop Squared*. These results are available from the authors on request.

Table 5. Change in Expected Number of Plants from a One Standard Deviation Change, Based on Coefficient Estimates from Specification 3, Table 4

Variable	% Change	Variable	% Change
High School %	30.99	*Occupation HHI*	-22.50
Union % (RTW = 0)	-8.69	*Per Capita Prop Tax*	-8.40
Union % (RTW = 1)	13.06	*Wages/Value Added*	-5.48
Population	27.19	*Population Density*	1.38
Manuf Empl %	7.45	*Unemployed*	-7.20
Hispanic %	9.00	*Per Capita Income*	-2.65
Black %	-1.04	*Manuf Estab*	16.92
Gini	8.22	*Value of Crops*	1.86
Education HHI	-13.93	*Value of Livestock*	6.86

Table 6. Factors Affecting Urban over Rural Location

Variable	% Advantage	Variable	% Advantage
High School %	24.35	*Occupation HHI*	-1.68
Union % (RTW = 0)	-3.46	*Per Capita Prop Tax*	0.86
Union % (RTW = 1)	5.21	*Wages/Value Added*	2.44
Population	26.98	*Population Density*	0.79
Manuf Empl %	0.76	*Unemployed*	2.43
Hispanic %	0.50	*Per Capita Income*	-0.85
Black %	-0.15	*Manuf Estab*	14.92
Gini	-5.83	*Value of Crops*	0.65
Education HHI	10.85	*Value of Livestock*	0.00

Agglomeration economies associated with population and prior manufacturing presence create substantial advantages for urban counties. Wages, which are typically lower in rural areas, do not seem to create an advantage for rural areas once adjusted for productivity. Finally, the supply variables, *Value of Crops* and *Value of Livestock*, apparently do not offset the disadvantages faced by rural counties in other variables.

Robustness Checks

Results from estimating our model for each of nine subsamples, defined by three-digit SIC code, are presented in tables 7a, 7b, and 7c. Results from each subsample usually agree with results from the overall sample. Notably, the signs for the heterogeneity coefficients agree with those in the overall sample, and are most of the time statistically significant. The only SIC industry that appears to be unaffected by labor-force heterogeneity is Fats and Oils, although the coefficients have the expected signs even in this subsample.

Some results are counter to our findings with regard to racial diversity. In particular, the percentage of a county's population that is Black (*Black %*) is statistically significant

Table 7a. Results of Model Estimation for Food Industry Subsamples: Meat, Dairy, and Preserved Fruits & Vegetables

	Meat		Dairy		Preserved Fruits & Vegetables	
Variable	Parameter	Std. Error	Parameter	Std. Error	Parameter	Std. Error
Constant	-0.579	1.025	-4.039**	1.737	-3.715***	1.367
Black %	0.003	0.003	-0.011*	0.006	-0.004	0.004
Hispanic %	-0.003	0.004	0.008	0.006	0.017***	0.004
Gini	0.025*	0.014	0.062***	0.023	0.041**	0.017
Education HHI	-0.351**	0.150	-0.353	0.243	-0.503**	0.201
Occupation HHI	-0.945***	0.267	-1.331**	0.524	-1.343***	0.438
Suburban	-0.143	0.090	-0.150	0.140	-0.280***	0.106
Rural	-0.592***	0.096	-0.630***	0.160	-0.868***	0.125
High School %	0.016***	0.006	0.040***	0.010	0.055***	0.008
Union %	-0.008	0.005	-0.029***	0.008	-0.018***	0.006
Population	0.006**	0.003	0.010**	0.004	0.009***	0.003
Pop Squared	-1.7E-05***	2.1E-06	-2.1E-05***	3.1E-06	-2.0E-05***	2.4E-06
Manuf Empl %	0.010**	0.005	0.005	0.008	0.008	0.006
Per Capita Prop Tax	-0.0003*	1.3E-04	-7.2E-04***	2.4E-04	-3.0E-04*	1.7E-04
Wages/Value Added	-0.126	0.333	-0.553	0.641	-1.257**	0.517
Population Density	4.2E-05***	1.3E-05	2.0E-05	2.1E-05	-3.4E-06	1.4E-05
Unemployed	-0.024	0.015	-0.074**	0.029	0.030	0.021
Per Capita Income	-0.015**	0.007	-0.001	0.004	-0.001	0.003
Highway	0.225***	0.065	0.371***	0.112	0.230***	0.088
Right-to-Work (RTW)	-0.454***	0.169	-0.577**	0.281	-0.781***	0.219
RTW × Union %	0.025**	0.011	0.033	0.020	0.045***	0.015
Manuf Estab	0.007***	0.002	0.007***	0.003	0.005***	0.002
Value of Crops	-0.004	0.005	0.000	0.007	0.009*	0.005
Value of Livestock	0.042***	0.006	0.027***	0.009	0.005	0.007
New England	-0.099	0.210	1.394***	0.295	0.545**	0.227
Mideast	-0.038	0.182	1.903***	0.282	0.624***	0.222
Great Lakes	0.143	0.161	1.326***	0.267	0.452**	0.206
Plains	0.325***	0.116	0.404*	0.212	-0.081	0.163
Southwest	0.087	0.127	0.040	0.238	-0.206	0.172
Rocky Mountain	0.455***	0.159	0.554**	0.271	0.081	0.208
Far West	-0.046	0.172	0.838***	0.269	0.488**	0.193
alpha (α)	0.339***	0.053	0.740***	0.125	0.351***	0.069
Log Likelihood	-2,835.25		-1,482.03		-1,882.42	
$R^2_{DEV,NB}$	0.266		0.415		0.443	

Note: Single, double, and triple asterisks (*) denote statistical significance at the 0.10, 0.05, and 0.01 levels, respectively.

Table 7b. Results of Model Estimation for Food Industry Subsamples: Grain Mill, Bakery, and Confections

	Grain Mill		Bakery		Confections	
Variable	Parameter	Std. Error	Parameter	Std. Error	Parameter	Std. Error
Constant	-3.150***	1.156	-4.222***	1.143	-3.721**	1.701
Black %	-0.002	0.004	-0.005	0.004	0.000	0.005
Hispanic %	-0.003	0.005	0.013***	0.004	0.001	0.006
Gini	0.061***	0.016	0.047***	0.014	0.064***	0.022
Education HHI	-0.447***	0.166	-0.548***	0.176	-0.970***	0.259
Occupation HHI	-1.241***	0.319	-1.047***	0.352	-1.406**	0.558
Suburban	-0.019	0.098	-0.279***	0.088	-0.138	0.129
Rural	-0.382***	0.107	-1.228***	0.113	-0.934***	0.160
High School %	0.033***	0.007	0.063***	0.007	0.046***	0.010
Union %	-0.019***	0.006	-0.011**	0.005	-0.017**	0.008
Population	0.008***	0.003	0.014***	0.002	0.007**	0.003
Pop Squared	-1.7E-05***	2.2E-06	-1.9E-05***	1.7E-06	-1.7E-05***	2.5E-06
Manuf Empl %	0.025***	0.005	0.001	0.006	0.016**	0.008
Per Capita Prop Tax	-2.6E-04*	1.5E-04	-7.5E-05	1.4E-04	-7.7E-05	2.1E-04
Wages/Value Added	-0.821**	0.400	-0.667	0.449	-1.055	0.643
Population Density	1.5E-05	1.5E-05	1.9E-05**	1.2E-05	-4.0E-06	1.6E-05
Unemployed	-0.046**	0.018	-0.016	0.020	0.000	0.027
Per Capita Income	-0.005	0.006	0.000	0.001	-0.003	0.007
Highway	0.285***	0.072	0.199**	0.079	0.363***	0.112
Right-to-Work (RTW)	-0.410**	0.188	-0.134	0.181	-0.439	0.267
RTW × Union %	0.029**	0.012	-0.002	0.013	0.031*	0.018
Manuf Estab	0.006***	0.002	0.004***	0.001	0.007***	0.002
Value of Crops	0.007	0.005	0.004	0.004	0.013**	0.005
Value of Livestock	0.035***	0.006	-0.004	0.006	0.004	0.008
New England	0.228	0.223	0.486***	0.180	0.373	0.286
Mideast	0.706***	0.201	0.698***	0.171	1.103***	0.264
Great Lakes	0.528***	0.181	0.080	0.171	0.560**	0.255
Plains	0.554***	0.133	-0.466***	0.136	-0.147	0.207
Southwest	0.161	0.145	-0.510***	0.149	0.073	0.203
Rocky Mountain	-0.077	0.202	-0.259	0.171	0.179	0.263
Far West	0.519***	0.183	0.052	0.159	0.366	0.248
alpha (α)	0.350***	0.057	0.247***	0.038	0.407***	0.094
Log Likelihood	-2,460.44		-2,159.36		-1,430.78	
$R^2_{DEV,NB}$	0.317		0.664		0.449	

Note: Single, double, and triple asterisks (*) denote statistical significance at the 0.10, 0.05, and 0.01 levels, respectively.

Table 7c. Results of Model Estimation for Food Industry Subsamples: Fats & Oils, Beverages, and Miscellaneous

	Fats & Oils		Beverages		Miscellaneous	
Variable	Parameter	Std. Error	Parameter	Std. Error	Parameter	Std. Error
Constant	-0.122	2.512	-3.556***	1.370	-1.422	1.039
Black %	0.020***	0.007	0.006	0.005	0.007**	0.003
Hispanic %	0.003	0.010	0.022***	0.005	0.021***	0.003
Gini	0.021	0.033	0.038**	0.017	0.031**	0.013
Education HHI	-0.389	0.393	-1.100***	0.215	-0.395***	0.150
Occupation HHI	-0.883	0.655	-1.290***	0.433	-1.536***	0.326
Suburban	-0.309	0.215	-0.233**	0.108	-0.094	0.083
Rural	-0.833***	0.236	-0.812***	0.130	-0.826***	0.095
High School %	0.001	0.015	0.069***	0.008	0.039***	0.006
Union %	-0.024	0.015	-0.005	0.006	-0.016***	0.005
Population	0.006	0.006	0.006**	0.003	0.107***	0.024
Pop Squared	-2.2E-05***	4.9E-06	-1.9E-05***	2.5E-06	-2.1E-04***	2.0E-05
Manuf Empl %	0.011	0.011	0.013**	0.007	0.004	0.005
Per Capita Prop Tax	-1.2E-04	3.2E-04	-2.1E-04	1.7E-04	-4.2E-04***	1.3E-04
Wages/Value Added	-1.697*	0.950	-0.135	0.503	-1.067***	0.373
Population Density	2.3E-05	2.7E-05	-7.2E-06	1.5E-05	-2.2E-05*	1.3E-05
Unemployed	-0.142***	0.045	-0.028	0.023	0.010	0.015
Per Capita Income	-0.017	0.017	-0.002	0.002	0.000	0.001
Highway	0.527***	0.165	0.174**	0.089	0.048	0.065
Right-to-Work (RTW)	-1.302***	0.455	-0.577***	0.221	-0.385**	0.165
RTW × Union %	0.095***	0.027	0.013	0.016	0.024**	0.011
Manuf Estab	0.008**	0.004	0.006***	0.002	0.007***	0.002
Value of Crops	-0.019	0.013	0.001	0.005	0.003	0.004
Value of Livestock	0.036***	0.013	-0.009	0.008	0.004	0.005
New England	-0.266	0.556	0.368	0.230	0.844***	0.179
Mideast	0.150	0.462	0.495**	0.218	0.735***	0.171
Great Lakes	0.484	0.417	-0.026	0.212	0.329**	0.162
Plains	0.788***	0.289	-0.432**	0.172	-0.207*	0.123
Southwest	0.029	0.320	-0.522***	0.190	-0.206	0.128
Rocky Mountain	-0.242	0.509	0.310	0.201	0.046	0.160
Far West	1.240***	0.392	0.889***	0.192	0.418***	0.152
alpha (α)	1.383***	0.325	0.490***	0.067	0.370***	0.047
Log Likelihood	-879.32		-1,923.34		-2,848.00	
$R^2_{DEV,NB}$	0.259		0.548		0.488	

Note: Single, double, and triple asterisks (*) denote statistical significance at the 0.10, 0.05, and 0.01 levels, respectively.

for Dairy, Fats and Oils, and Miscellaneous, and *Hispanic %* is statistically significant for Preserved Fruits and Vegetables, Bakery, Beverages, and Miscellaneous.[14]

Conclusion

We have used a negative binomial regression to model food manufacturing plant entry and to analyze the importance of labor-force heterogeneity for attracting new food plants. Our model fits the data well, and many coefficient estimates are in accord with previous studies. Large and statistically significant effects emanate from various measures of labor-force heterogeneity. This result is consistent with findings by Henderson, Kuncoro, and Turner (1995) and Duranton and Puga (2001), who report empirical evidence that dynamic externalities, such as labor-force heterogeneity, foster growth in cities. However, these authors found the externalities to be important for young industries. In contrast, our analysis is conducted on a mature industry—food manufacturing—and yet it finds strong evidence in favor of these externalities.

Previous research has included an area's racial characteristics as a potential determinant of establishment entry. As in some of this previous research, we also find a positive and statistically significant relationship between plant entry and the proportion of a county's population that is Black, when not controlling for labor-force heterogeneity. However, when measures of labor-force heterogeneity are included in a model with *Black %*, this variable loses statistical significance. This finding suggests that variables measuring racial diversity (like *Black %*) may be capturing the effect of labor-force heterogeneity in these earlier studies, and it may be labor-force heterogeneity that is important to plant managers, not the racial make-up of an area. This result is somewhat mitigated because the percentage of Hispanic workers (*Hispanic %*) is significant in model specifications 1–3, and in many of the subsample estimations. Moreover, checking the robustness of the *Black %* variable, it is significant in the Dairy, Fats and Oils, and Miscellaneous subsamples. Nonetheless, we believe the results of this analysis raise interesting and important questions given the strong relationship between a county's heterogeneous labor force and the likelihood of new food plant entry.

This result provides guidance for policy makers whose goals are to attract food manufacturing to their locality. The quality of an area's labor force is an important feature considered by investors when choosing a location for a new food plant. However, labor force quality is multi-dimensional and localities that have workers with a variety of skills have an advantage over localities with worker skills concentrated in a few areas.

Finally, we note that our recommendations rely on a properly specified model. We contend our results suggest a causal relationship running from labor-force heterogeneity to plant location decisions. An alternative explanation is that plant clustering *prior* to

[14] We also experimented with methods to gravity weight key independent variables (e.g., market and supply variables) to reflect the influence of neighboring counties. Variables were weighted by the inverse of the county-to-county distance between county centers. For own-county weight, we used the inverse of the county radius. Because it is not possible to nest the gravity model within the nongravity model (or vice versa), we attempted to discern the appropriate model with nonnested tests. In nonnested tests, either model can serve as the null, and unfortunately in our case when each model served as the null it was rejected in favor of the alternative. Without this statistical guidance, we resorted to basing our judgment on the Bayesian information criterion (BIC), or equivalently the model with the superior log-likelihood value. The reported model with unweighted variables provided a better fit based on a BIC (Cameron and Trivedi, 1998), and key results from the weighted model were qualitatively consistent with those reported.

plant entry attracts a heterogeneous workforce, just as clustering generates agglomeration economies and attracts new food plants—in which case, the heterogeneity variables may only be a proxy for agglomeration economies, rather than a source for agglomeration economies. However, if we have properly controlled for agglomeration economies in our specification, then our results and policy recommendations hold.

[*Received February 2005; final revision received November 2005.*]

References

Adhikaril, B. B., S. B. Harsh, and L. M. Cheney. "Factors Affecting Regional Shifts of U.S. Pork Production." Selected paper presented at annual meetings of the American Agricultural Economics Association, Montreal, Canada, July 27–30, 2003.

Alesina, A., R. Baqir, and W. Easterly. "Public Goods and Ethnic Divisions." *Quart. J. Econ.* 114,4 (November 1999):1243–1284.

Bartik, T. "Business Location Decisions in the United States: Estimates of the Effects of Unionization, Taxes, and Other Characteristics of States." *J. Bus. and Econ. Statis.* 3,1(January 1985):14–22.

Cameron, A. C., and R. K. Trivedi. *Regression Analysis of Count Data.* London/New York: Cambridge University Press, 1998.

Cameron, A. C., and F. G. Windmeijer. "R-squared Measures for Count Data Regression Models with Applications to Heath-Care Utilization." *J. Bus. and Econ. Statis.* 14(April 1996):209–220.

Chappell, W. F., M. S. Kimenyi, and W. J. Mayer. "A Poisson, Probability Model of Entry and Market Structure with an Application to U.S. Industries During 1972–77." *S. Econ. J.* 56(April 1990): 918–927.

Coughlin, C. C., and E. Segev. "Location Determinants of New Foreign-Owned Manufacturing Plants." *J. Regional Sci.* 40(2000):323–351.

Duetsch, L. L. "Entry and the Extent of Multiplant Operations." *J. Industrial Econ.* 32(June 1984): 477–488.

Dunne, T., M. J. Roberts, and L. Samuelson. "Patterns of Firm Entry and Exit in U.S. Manufacturing Industries." *Rand J. Econ.* 19(Winter 1988):495–515.

Duranton, G., and D. Puga. "Nursery Cities: Urban Diversity, Process Innovation, and the Life Cycle of Products." *Amer. Econ. Rev.* 91,5(December 2001):1454–1477.

Friedman, J., D. A. Gerlowski, and J. Silberman. "What Attracts Foreign Multinational Corporations? Evidence from Branch Plant Location in the United States." *J. Regional Sci.* 32(November 1992): 403–418.

Glaeser, E. L., H. D. Kallal, J. A. Scheinkman, and A. Shleifer. "Growth in Cities." *J. Polit. Econ.* 100 (1992):1126–1152.

Goetz, S. J. "State- and County-Level Determinants of Food Manufacturing Establishment Growth: 1987–93." *Amer. J. Agr. Econ.* 79(August 2000):838–850.

Guimarães, P., O. Figueiredo, and D. Woodward. "A Tractable Approach to the Firm Location Decision Problem." *Rev. Econ. and Statis.* 85(February 2003):201–204.

———. "Industrial Location Modeling: Extending the Random Utility Framework." *J. Regional Sci.* 44(February 2004):1–20.

Hausman, J. A. "Specification Tests in Econometrics." *Econometrica* 46(November 1978):1251–1271.

Henderson, J. R., and K. T. McNamara. "The Location of Food Manufacturing Plant Investments in Corn Belt Counties." *J. Agr. and Resour. Econ.* 25(December 2000):680–697.

Henderson, V., A. Kuncoro, and M. Turner. "Industrial Development in Cities." *J. Polit. Econ.* 103 (1995):1067–1090.

Hogler, R. L. *Labor and Employment Relations.* St. Paul, MN: West Publishing Co., 1995.

Hopkins, J. "Fatality Rates Increase for Hispanic Workers." *USA Today* (13 March 2003).

Jacobs, J. *The Economies of Cities.* New York: Random House, 1969.

Kim, S. "Labor Heterogeneity, Wage Bargaining, and Agglomeration Economies." *J. Urban Econ.* 28(1990):160–177.

List, J. A. "U.S. County-Level Determinants of Inbound FDI: Evidence from a Two-Step Modified Count Data Model." *Internat. J. Industrial Org.* 19(2001):953–972.

List, J. A., and M. Kunce. "Environmental Protection and Economic Growth: What Do the Residuals Tell Us?" *Land Econ.* 76,2(May 2000):267–282.

MacDonald, J. M., M. Ollinger, K. Nelson, and C. Handy. "Consolidation in U.S. Meatpacking." Agr. Econ. Rep. No. 785, USDA/Economic Research Service, Washington, DC, 2000.

Orr, D. "The Determinants of Entry: A Study of the Canadian Manufacturing Industries." *Rev. Econ. and Statis.* 56(February 1974):58–66.

Schmenner, R. W., J. C. Huber, and R. L. Cook. "Geographic Differences and the Location of New Manufacturing Facilities." *J. Urban Econ.* 21(1987):83–104.

Smith, D. F., and R. Florida. "Agglomeration and Industrial Location: An Econometric Analysis of Japanese-Affiliated Manufacturing Establishments in Automotive-Related Industries." *J. Urban Econ.* 36(1994):23–41.

U.S. Department of Commerce, Bureau of the Census. *Statistical Abstract of the United States: 2004–2005*, 124th edition. Washington DC: Government Printing Office, 2004. Online. Available at http://www.census.gov/statab/www/.

Woodward, D. P. "Locational Determinants of Japanese Manufacturing Start-ups in the United States." *S. Econ. J.* 58(January 1992):690–708.