PERE
7/26/06
1:00 PM

Networking Off Madison Avenue[1]

Mohammad Arzaghi and J. Vernon Henderson

Brown University

July, 2006

New York County, or Manhattan, accounts for 24% of advertising agency receipts in the USA. Years ago the popular image had large agencies on Madison Avenue generating this business. But today there are over 1000 agencies in different clusters spread over different neighborhoods in the southern half of Manhattan. In clustering, advertising is known for the key role that networking plays in the operation of agencies, where interviews suggest that informal networking with exchange of ideas and expertise among close neighbors is critical to many agencies' success (Arzaghi, 2005). We interpret localized networking as a structured case of information spillovers, or Marshall's (1890) "mysteries" of the "air", involving more deliberate information exchanges.

This paper examines the effect on productivity of having more near advertising agency neighbors and hence better opportunities for meetings and information exchange within Manhattan. We have census tract data for Manhattan that allows us to distinguish locations at 250 meter increments. With this fine level of geographic detail, we can infer the high benefits of close social interactions. We show that there is extremely rapid spatial decay in the benefits of having more near neighbors even in the close quarters of southern Manhattan, a finding that is new to the empirical literature. This suggests that having a high density of similar commercial establishments is important in enhancing local productivity for certain industries, as modeled in Lucas and Rossi-Hansberg (2002).

Our results indicate that, in Manhattan, advertising agencies trade-off the higher rent costs of being in bigger clusters closer to the "centers of action", against the lower rent costs of operating on the "fringes" away from high concentrations of other agencies. That trade-off indicates that the benefits of higher spillovers, or better opportunities to network, within metropolitan areas such as Manhattan, may be largely capitalized into rents, rather than wages. .

---

[1] The research in this paper was conducted while the authors were Census Bureau research associates at the Boston Research Data Center. Research results and conclusions expressed are those of the authors and do not necessarily indicate concurrence by the Census Bureau. The paper has been screened to ensure that no confidential data are revealed. We acknowledge helpful comments from Andrew Foster on aspects of the paper and from participants in seminars at Brown University, LSE, and the University of Toronto.

This study differs in key ways from traditional work on urban scale externalities and challenges the existing methodology for examining these externalities. First, almost the entire literature, as reviewed in Rosenthal and Strange (2004), examines externalities in manufacturing. Manufacturing to the extent it still plays an important role in American city productivity is found disproportionately in rural areas and smaller and medium size cities (Kolko, 1999). If we want to understand the *raison d'être* of larger metropolitan areas and why certain industries are drawn to high rent, high wage cities like New York with historically built dense neighborhoods, we need to look at high-end business and financial services, such as advertising services.

Second, the literature on externalities examines either total factor productivity benefits (e.g., Henderson, 2003) or wage gains (e.g., Glaeser and Mare, Rosenthal and Strange, 2005) of greater agglomeration at the aggregate county or metropolitan area level. One problem just noted is that, for certain service industries at least, localized information spillovers are capitalized into commercial rents, not wages. The other is that key aspects of information spillovers are captured by within county, not across county variation. For advertising agencies, networking benefits are so localized, that they are exhausted within the confines of sub-areas of southern Manhattan: firms on the far west side of southern Manhattan derive no direct benefits from firms on the east side. While Rosenthal and Strange (2003) introduce spatial decay into externality estimation, they look just at manufacturing (and software) and assume no spatial decay within the first mile, while here effects completely decay within a half mile!

Third, this literature debates what the nature of scale benefits is—information spillovers, labor market externalities, or reduced transactions and transport costs of exchange of intermediate inputs as in the new economic geography (Fujita, Krugman, and Venables, 2001)—with no clear resolution. Looking within a county allows us to isolate specific sources of externalities. We know what we capture are not labor market externalities: Southern Manhattan is all one labor market, drawing workers from all over the New York PMSA and even beyond.[2] We will argue below that finding immediate spatial proximity to other advertising agencies matters implies proximity benefits for information exchange and networking, rather than some type of shopping-center externality (Dudey, 1990) or market potential effect. If what we isolate within Manhattan involves information spillovers, that raises the issue of how to interpret cross

[2] There is evidence of very modest wage gradients (over longer distances than just within Manhattan) in large cities; but it is ascribed to the fact that those living further from the center earn lower wages because they work locally and have lower commuting costs, rather than there being different labor markets (McMillan and Singell, 1992).

county or metropolitan area results which find aggregate scale of own industry activity matters. For advertising agencies, they could represent labor market or shopping-center externalities at the metropolitan area level; or they could represent greater opportunities for clustering and networking in larger metropolitan areas.

In terms of scale externalities outside the own industry, we explore the benefits of immediate access to buyers (e.g., headquarters), suppliers (e.g., graphics) and outlets (broadcasters). Only access to broadcasters matters, which represents the new economic geography costs of trade (negotiating and buying advertising time). But immediate proximity to headquarters or graphic services does not matter. The latter are ubiquitous and the former are a short cab ride away in Manhattan where day-to-day, face-to-face interaction is not required. In fact, as discussed below, most advertising services produced in Manhattan are now exported to companies away from New York City.

Finally, another advantage of looking within counties is that cross MSA and county studies are plagued by unmeasured characteristics which affect both productivity and local industry scale and vary across large geographic units. These include legal, regulatory and business climate, access to specific markets, and specific infrastructure including the transport system which affects commuting and work effort. In a panel context one can try to control for these with fixed effects (Henderson, 2003), but such effects may not be time invariant, especially over longer panels. Southern Manhattan is all one legal-regulatory-business climate and infrastructure system, so the usual cross-section unobservables are not an issue in identification. There are tract level unobservables; but we argue they are easier to deal with econometrically, so that identification is more straightforward than cross county or MSA level analysis.

1 Advertising Agencies

In this section we give some background information on the advertising agency industry, in general and in Manhattan. Then we look at Manhattan in more detail.

1.1 General Background

In 1997, advertising expenditures in the U.S. totaled $187b, about 2.2% of GDP. Of this, only $17b involves advertising agency receipts (NIPA, BEA).[3] However of the $17b, $10b comes from commissions on $89b of media billings carried out by advertising agencies. This implies that over half of advertising expenditures in the U.S. go through advertising agencies and reflects the fact that many advertising agencies continue the historical function of intermediation, or matchmaking between advertisers and media outlets. However, advertising agencies are more popularly known for the design of advertising campaigns. From the industry point of view, top notch firms provide "great creative thinking campaigns and ideas" for advertisers to market their products and to develop a "unique selling proposition" (Schemetter, 2003).

Advertising agencies operate with internal teams, where a team typically provides the full range of services for a client—market research, creation of an advertising campaign, production, media, client management, and the like, where most agency costs are labor costs. Apart from potential media commissions, advertising agencies derive a significant portion of their income from fees and billings for creative services rendered in the design of campaigns.

To win an account, an agency team competes by developing new ideas. Thus, like R&D (Jaffe et al., 1993), with creativity playing a central role, information sharing and information diffusion are thought to be critical to an agency's success. That sharing occurs within and across agencies, where agencies belong to formal and informal networks.[4] Based on interviews with agencies in New York and Chicago, informal networking appears to occur at the design stage of advertising campaigns where an agency has received a "request for proposal" (or RFP) for an advertising campaign (Arzaghi, 2004a). The agency turns to members of its local network for information (who by odds are unlikely to have received the same RFP) on how best to respond to the request for proposal and how to design the potential campaign. In New York, we believe clustering occurs because of localized networking to enhance creativity; agencies share information and ideas where repeated face-to-face contact is critical. Contacts may be "official"

---

[3] Advertising is split about 55-45 between national versus local advertising (Berndt and Silk, 1994) and is focused on automobiles (18%), retail, department and discount stores (15%) and then movies, cosmetics and toiletries, medicines, food, financial services and restaurants with shares each of 4.3-5%. In terms of media, TV, radio and cable TV account for 36%, magazines and newspapers 32% and yellow pages and direct mail 30%. The last two sets of numbers are from the AdAge website, based on reports by respectively TNS Media Intelligence and Coen/McCann-Erikson. Both data sets are property of Crain Communication, Inc..

[4] Data exist on the formal network structure of advertising agencies, where members pay dues to networks which have full-time managers and they attend network meetings several times a year. Hameroff (1998) lists eleven huge national and international formal networks of larger agencies. These formal networks primarily exchange information about marketing conditions across cities and countries.

such as meetings among members of a neighborhood network, unofficial such as informal contacts in local lunch and coffee places, or both.

The notion that networking is so important to advertising agencies compared to most other activities may be related to the fact that advertising agencies have low industrial concentration, in terms of both establishment and firm size. Table 1 compares advertising agencies relative with business services more generally, noting business services are much less concentrated than manufacturing. Establishments under five employees account for 58% of all establishments and 14% of sales, higher than for business services generally. The average establishment size is 11 employees compared to about 50 in manufacturing and to 21 in the rest of business services. Single-unit firms account for 90% of establishments and 55% of sales in the advertising agency industry, compared to 81% and 45% in other business services. At the upper tail, the 17 largest (in terms of number of establishments) multi-unit firms only account for 22% of industry sales.

While the advertising agency industry is a non-concentrated industry, it is heavily spatially concentrated in New York County, or Manhattan, as shown in Table 2. In 1997, Manhattan has 7.3% of advertising agency establishments in the U.S., 20% of their employment, 24% of all advertising agency receipts, and 31% of media billings. For comparison, we note that Manhattan has about 1.9% of all private employment in the USA; so location quotients for New York for advertising agency activities are very high. The relatively fewer numbers of establishments (7.3%) compared to employment (20%) means establishment sizes are much larger in Manhattan than elsewhere. The high concentration of advertising agencies in Manhattan means these agencies "export" much of their product; their buyers are outside Manhattan (noting also that Manhattan's share of headquarters has declined over time, now garnering only 3% of national headquarters employment).

In the empirical work, in examining the advertising agency business, we focus on the activities of single-unit [SU] firms, as opposed to multi-unit [MU] firms. Why don't we pool the groups? In the text and Appendix B, we will argue that, in econometric work, firms in each group seem to interact within the own group (SU and MU), but not across groups. MU and SU's have very different establishment sizes and types of activities. In New York, SU's have an average size of 11 employees, while MU *establishments* (not firms) have an average size of 155. MU's are much more reliant on income from media receipts indicating their greater role as intermediaries between advertisers and the media, In 1997 in NY, they got 48% of receipts from media billings, while SU's got 28%. In contrast, SU's got 21% of their income on commissions and fees for services provided by *other* agencies, compared to just 6% for MU's, an indication of greater "sharing" and cross-firm indirect sales by SU's. And most of the rest of SU income is from fees for their own services. In general, SU's are the firms where localized cross-firm networking is widespread and close spatial clustering is important. In contrast, for their more media based activities, MU's can rely on intra-firm networks across their own establishments. Thus we separate MU's from SU's and focus on SU's with their large sample size.

1.2 Manhattan

Within Manhattan, advertising agencies are closely clustered. Because of disclosure criteria, we are not allowed to map agency locations using Census data. Although the empirical work to follow is based on the entire population of advertising agencies in the Census Business Register (or Standard Statistical Establishment List—SSEL), we use Advertising Red Books online for 2003 to generate maps. The Red Books give the location of about 800 of the 1000 or so agencies in Manhattan.[5] Figure 1 shows advertising agency locations in the southern half of Manhattan, defined as the area below 90th Street. Almost all agencies in NY are located below 90th Street. Overall, clustering is heaviest within a block or two of Madison and Fifth Avenues, but there appear to be many "sub-clusters" as we move north-south along the Avenues, looking both east and west. The Red Books do not reliably breakdown single and multi-unit firms, so we can't map them separately. If we compare Figure 1 with the rent gradient map in Figure A-2, we can see there are large clusters of agencies in medium rent areas, with somewhat fewer in the highest rent areas. In low rent areas on the fringes, there are even fewer agencies. For 1992, the simple correlation between the number of SU's in a tract and the log of rent price is .23.

In order to further explore the spatial distribution of agencies within Manhattan, we return to Census data. We examine similarity of location patterns by zip code, asking whether any comparison pair (e.g., SU stocks in 1977 versus SU stocks in 1992, or SU vs. MU stocks in 1992) is drawn from the same distribution (across zip codes). We use zip code location information here, rather than the Census tracts we use later, since only the former are available in

[5] This is a publication of National Register Publishing. The Red Book includes self reported data.

all Census comparison years (1977, 82, 87, 92, 97). First, Pearson χ^2 tests[6] show MU and SU establishment stocks have different location patterns across zip codes in all years (1977, 82, 87, 92, 97), consistent with the hypothesized lack of SU-MU interaction noted earlier.[7]

Next we examine a key aspect of SU location patterns, relevant in the sections to follow, relying on χ^2 tests to compare location patterns across tracts for all of Manhattan. The key result is that location patterns of SU's have changed overtime, which is useful information in discussing identification later.[8] Specifically:

(1) Comparing 1987 (or 1982) with 1997, the SU stock location pattern in 1987 differs from the SU pattern in 1997.

(2) Corresponding to (1), births of SU's in 1982-87 vs. 1992-97 have different spatial distributions.

Patterns change from 1987 to 1997 reputedly in response to relative rental cost changes at that time, moving advertising agencies out of some of the highest rent areas. This effect of changing rents is thought to be enhanced by an "exogenous" change in revenue structure for agencies. Until the early 1980's, agencies received a 15 percent commission rate on media purchases they made for advertisers. In part following through on a Justice Department decree, this rate was abandoned in the earlier 80's and by 1997 had fallen to 11%. This supposedly induced cost cutting measures (e.g., moving to lower rent districts), as well as helped refocus activities of some agencies. Most births since 1992 have happened in medium and lower rent areas and hence off Madison Avenue.

2. A Model of Expected Networking and Location Choice

In this section, we develop a model of location choice within New York City for single-unit advertising agencies, to help us think about empirical issues. We start with the specification

---

[6] The test statistic is $T = \sum_{i=1}^{r}\sum_{j=1}^{c}\frac{(m_{ij}-e_{ij})^2}{e_{ij}}$ where $e_{ij} = Np_i p_j \equiv N\frac{n_{i\cdot}}{N}\frac{n_{\cdot j}}{N}$ is the ML estimate of probability of cell ij where $n_{i\cdot}$ is sum across rows (say, SU and MU) and $n_{\cdot j}$ is sum across columns (say, locations), and m_{ij} is the count in any cell. T is distributed χ^2 with degree of freedom (r-1) (c-1).

[7] We cannot reject, for each of the periods 92-97, 87-92, and 82-87, that SU and MU births have the same location patterns as each other. However for both of 87-92 and 92-97, the p-values are 0.052, on the border of rejection of similarity.

[8] These results on the first three items also apply (separately) to MU's.

of networking technology and location choice. In section 3, we turn to the details of implementation, empirical specification, data, estimation, and results. Section 4 provides conclusions and analyzes some extensions.

2.1 Formulating Local Network Contacts

We specify a simple model for a new firm that is choosing a location in anticipation of informally networking with other agencies nearby. The distance between agencies plays a key role in the extent of networking. While the specification and jargon is informed by the large theoretical literature on networks (e.g., Jackson and Wolinsky, 1996, Bala and Goyal, 2000, and Goyal and Moraga-Gonzalez, 2002), we look at just one firm and how to specify the benefits, costs and extent of networking for it. We model networking as overlapping with direct linkages only, where each pair of establishments decides on its pair-wise extent of contact, independent of other pair-wise arrangements.

The idea is that a new firm anticipates networking with advertising agencies in close-by locations, where this networking can involve social "meetings" with employees of other agencies (e.g., going to the bar after work or coffee in the morning), scheduled business meetings where some team members in a firm get together with some team members in another firm, or contacts through intermediaries such as local graphics firms or contract designers serving several agencies. A new firm in networking with any other firm has a match parameter, revealed after the firm has picked its location and established contact with the other firm. Given the *ex post* chosen level of networking for each pair, networking is costly in terms of time, where that cost rises with distance between pairs. This could represent commuting time for the requisite face-to-face meetings or monitoring of one another's activities.

Although our estimation is not strictly structural, we provide a framework to motivate our specification and interpret results for those who want a more structural interpretation. Consider a firm operating at location j. The value of its contact with a firm at location i is

$$A_j[V_{ij}^{\gamma}\ u_{ij} - \ V_{ij}\mathrm{c}(\delta_{ij})] - F\ . \qquad (1)$$

In (1), V_{ij} is the volume of communications j gets from a firm at i and $0 < \gamma < 1,$ so there are decreasing returns to volume in any pair of contacts. u_{ij} is a "match" parameter drawn on the interval [0,1], which is firm specific between the firm at j and a firm at i, and gives variation in the extent of contacts across firms. $c(\delta_{ij})$ is the cost per unit of communication, where δ_{ij} is the distance from i to j. Total communication costs comprise firm labor costs; one can define V_{ij} as labor units involved in communication, with $c(\delta_{ij})$ representing lost work time from employees having to travel further to interact with another firm. Differences in A_j could represent differences across locations in amenities. F is the initial fixed cost of making a contact with a firm at another location, upon which u_{ij} is learned.[9]

In formulating (1) and in interpreting results, as noted above, advertising agencies mostly sell to firms outside Manhattan and initial "shopping" of advertisers for advertising agencies is done by mail, telephone, and internet, not by "walk-in-traffic". Thus, unlike a retail concern, we don't expect market potential effects or shopping-center externalities within Manhattan. Nonetheless, in the empirical formulation we will examine and discuss formulations that allow for shopping-center externalities and/or market potential effects.

In our formulation, once the firm spends F to learn u_{ij}, it chooses V_{ij} so that $V_{ij}^* = (\gamma/\mathrm{c}(\delta_{\mathrm{ij}}))^{1/1-\gamma} \ (u_{ij})^{1/1-\gamma}$. If $c(\delta_{ij}) = c(\delta_{ji})$ and $u_{ij} = u_{ji}$, then in this simple case $V_{ij} = V_{ji}$ and communications are reciprocal and two sided. In the present specification, *ex post* profit from contact with any firm is $A_j(1-\gamma)\gamma^{\gamma/(1-\gamma)}c(\delta_{ij})^{-\gamma(1-\gamma)}(u_{ij})^{1/(1-\gamma)} - F$ and *ex ante* profit is

$$A_j(1-\gamma)\gamma^{\gamma/(1-\gamma)}c(\delta_{ij})^{-\gamma(1-\gamma)}E[u^{1/(1-\gamma)}] - F \qquad (2),$$

where we denote the expected value of $(u_{ij})^{1/(1-\gamma)}$ as $E[u^{1/(1-\gamma)}]$, which we assume is common across all i, for now. Note the expected volumes of communications and profits are decreasing in distance and communication costs.

Equation (2) relates to the key set of coefficients we will attempt to estimate—how expected profit to a firm at *j* is increased if a firm is added at *i* and how that profit declines with distance, δ_{ij}. Equation (2) also can be used to define the maximum distance, $\delta_{\max}$, over which a

---

[9] While u_{ij} can take the value zero to represent "no match" possibility, if we want to incorporate issues of conflict of interest across agencies, we would want to have a probability p of any contact resulting in "no match" and formulate an expected fixed cost per contact as $F/(1-p)$, with only actual retained contacts contributing to variable profits.

firm makes contacts and scale externalities are relevant. Empirically we will see that there is such a cut-off point. Given the positive term in (2) is declining in δ_{ij}, $\delta_{\max}$ is defined by

$$\delta_{\max} = \max\{\delta_{ij} \mid A_j (1-\gamma)\gamma^{\gamma/(1-\gamma)} c(\delta_{ij})^{-\gamma(1-\gamma)} E[u^{1/(1-\gamma)}] - F > 0\}. \qquad (3)$$

$\delta_{\max}$ is increasing in firm productivity, A_j, and in expected match productivity $E[u^{1/(1-\gamma)}]$, and is declining in costs.

In this simple context, total firm profits are simply the value of all contacts the firm makes minus the costs. Thus total expected profits are

$$\mathrm{E}[\Pi_j] = A_j \gamma^{\frac{\gamma}{1-\gamma}} (1-\gamma)\ \mathrm{E}[u^{\frac{1}{1-\gamma}}] \sum_{i}^{\delta_{\max}} c(\delta_{ij})^{-\frac{\gamma}{1-\gamma}} - FN(\delta_{\max}) - R_j\,, \qquad (4)$$

where R_j is the rental cost at location j and $N(\delta_{\max})$ is the number of agencies within the distance $\delta_{\max}$ of location j. Given the definition of $\delta_{\max}$ in (3), it is clear that having more firms within $\delta_{\max}$ (i.e., greater the density) raises profits. This presents two issues: why is there more than one cluster in New York and how might we start to think about heterogeneity of firms?

<u>Multiple clusters</u>. Why are there multiple clusters in Manhattan? First, in the empirics, the strategy is to have a firm-location specific matching parameter drawn in the usual logit fashion, so firms have preferences over locations per se, and thus spread out, some in distinct clusters and some more in isolation. Second, in Manhattan the number of available spaces within a neighborhood is limited and it isn't clear there could be a neighborhood of 1000 agencies. Third, one could specify the model so that there are decreasing returns to scale to a firm in overall volume of all communications limiting the extent of contacts; Arzaghi (2004) models optimal network sizes in competing for ad agency accounts, even without spatial decay. Finally, another motivation for multiple clusters is heterogeneity, the empirics of we discuss in the last section.

<u>Heterogeneity.</u> A context where there is variation in what different quality firms are willing to pay for features of advertising agency clusters has the potential to deliver a separating equilibrium as in Arzaghi (2005). There, high quality firms cluster in high cost neighborhoods and lower quality ones operate on the fringes because they can't afford the higher costs in larger clusters. Differences in firm quality could be reflected by differences across firms in γ or in the specification of matching parameters. For the latter we could assume each firm is endowed with its own "communication ability" parameter u_j, a measure of firm quality. For any pair, the match gets $u_{ij} = \min[u_j, u_i\ ]$, where u_j is the communication parameter for our firm at j and u_i is the parameter for a specific firm at i. Given $u \in [0,1]$, one can show that

$$E[(u)^{1/(1-\gamma)}] = (u_j)^{1/(1-\gamma)} - \int_0^{u_j} \frac{1}{1-\gamma} u^{\gamma/(1-\gamma)} G(u)du\,, \qquad (5)$$

where $G(u)$ is the distribution function. Equation (5) suggests higher quality firms have higher $E[(u)^{1/(1-\gamma)}]$; and they benefit more from neighbors in equations (2) or (4), so that their scale effects are enhanced. They also benefit more from being in a neighborhood with better quality neighbors, in the sense of a better distribution function, $G(u)$.[10] Higher quality firms will have more profits and be larger (more volume of communications) and thus in our interpretation will have larger employment. Later in the paper, we will use size to measure firm quality..

2.2 Empirical Implementation and Data

The model we estimate is based on equation (4). In the treatment of space, locations are census tracts, which in NY are at a finer level of spatial detail than zip codes. Moreover, a census tract is intended to represent a true micro "neighborhood," as most tracts in Manhattan coincide with the city blocks between two avenues and a few streets. Each census tract has a rental cost per square foot of class A office space, availability of commercial sites measured by the stock of private establishments, and distance measures to other advertising agencies and other activities such as broadcasting establishments (where ads are placed), headquarters (clients), or graphic services (inputs).

In the empirical implementation of equation (4) we want to allow office space sizes to respond to price, rather than being fixed. We adjust our specification of production technology, to yield a simple estimating equation. If l_j is the amount of space consumed at a unit rental cost, R_j, suppose *ex post* profits are specified as

[10] If a firm chooses a neighborhood with a better distribution, $F(u)$, of neighbors in the sense of first degree stochastic dominance, then $E[(u)^{1/(1-\gamma)}]$ rises by $\int_0^{u_j} \frac{1}{1-\gamma} u^{\gamma/(1-\gamma)} (G(u) - F(u))du > 0$. This rise in $E[(u)^{1/(1-\gamma)}]$ is greater for higher quality firms by $(1-\gamma)^{-1}(u_j)^{\gamma/(1-\gamma)}(G(u_j) - F(u_j))$.

$$A_j l_j^{\alpha} \exp\left(\sum_{i}^{\delta_{\max}} \{[V_{ij}^{\gamma} u_{ij} - \ V_{ij} \mathrm{c}(\delta_{ij})] - F\} \right) - R_j l \ .$$

Optimizing with respect to choice of V_j gives

$$A_j l_j^{\alpha} \exp\left(\sum_{i}^{\delta_{\max}} \{(1-\gamma)\gamma^{\gamma/(1-\gamma)} c(\delta_{ij})^{-\gamma(1-\gamma)} u_{ij}^{1/(1-\gamma)} - F\} \right) - R_j l_j \, .$$

For the choice of l_j, we take expected values, maximize *ex ante* profits, and substitute back in for the optimal value of l_j. Moreover, given the discrete nature of our geographical units (tracts), we allow there to be n_i firms at each location (tract) and then take the log of expected profits. After rearrangement, expected profits from choosing location *j* are

$$\tilde{\pi}_{\mathrm{j}} \equiv \ln \Pi_j = C_j - \frac{\alpha}{1-\alpha} \ln R_j + \frac{1}{1-\alpha} \sum_{i}^{\delta_{\max}} n_i \ln \mathrm{E}[\exp\{(1-\gamma)\gamma^{\frac{\gamma}{1-\gamma}} c(\delta_{ij})^{\frac{-\gamma}{1-\gamma}} u^{\frac{1}{1-\gamma}} - F\}] \ . \quad (6)$$

The empirics focus on estimating the effect of an increase in the number of firms at location *i* on the profits of a firm located at *j*, or a set of coefficients that decline with distance δ_{ij}, and match the expression $\ln \mathrm{E}[\exp\{(1-\gamma)\gamma^{\gamma/1-\gamma} c(\delta_{ij})^{-\gamma/1-\gamma} u^{1/1-\gamma} - F]$ in (6). Our estimation is not structural: to make it structural, we would need to assume a specific PDF for u_{ij} and shape of $c(\delta_{ij})$, for which we don't have good priors. Moreover the model is meant to motivate the empirical formulation and interpretation, rather than insist on the exact structure. Finally in (6), C_j includes arguments in A_j such as a firm-tract specific error term.

Formulation. In estimation we look at the location choices of births, or new firms in Manhattan, where firms pick the census tract which maximizes profits. In viewing location choices across tracts, as usual, we assume that the firm looks at each tract as "one" choice with a tract-firm idiosyncratic match parameter, η_j, drawn from an extreme value type I distribution. Tracts are defined to be neighborhoods in New York.

In a logit formulation the probability that a firm chooses tract j is

$$P_j = \frac{\exp(\pi_j)}{\sum_i \exp(\pi_i)} . \qquad (7)$$

In (7), π_j is $\tilde{\pi}_j$ in (6) without the C_j which includes the separable error term, η_j [11].

We augment the arguments of π_j by adding the log of the total number of private establishments in the neighborhood, m_i, for two reasons. First, it captures localized general agglomeration economies from having more commercial activities in a firm's own neighborhood, including any general market potential or shopping-center externalities beyond the specifics of headquarter locations in Manhattan which are treated separately. However, while we will find that the total number of establishments just within the own neighborhood, j, affects profits, these benefits, interestingly, do not extend even to neighbors in immediately adjacent areas as they would with usual market potential considerations in an economic geography framework. Finding that the total number of private establishments in only the own neighborhood matters is consistent with a different interpretation to own neighborhood total commercial scale suggested by Holmes (2005). Within the own tract, unobserved to us, spaces differ and rather than assuming the same error draw for a firm within a tract, we could assume each available location, z, in tract j is associated with its own idiosyncratic matching parameter to the firm, ε_{jz}, drawn from a generalized extreme value distribution. Office spaces on the same city-block in New York may be somewhat similar (in the same building) and certainly face similar street conditions, so it is natural to assume that these error drawings are correlated within tracts. A feature of Manhattan is that tracts differ in available spaces, where land and building usage is tightly regulated and relatively inflexible over time (Glaeser and Gyourko, 2002). If (unobserved) vacancies are proportional to existing commercial spaces and there are more available commercial spaces in a tract, then a firm enjoys more idiosyncratic draws and better matching possibilities within a tract. In that case, the coefficient of m_j is $1-\sigma$, where σ is approximately the within tract correlation of error drawings in a nested logit framework. Under this interpretation, equation (7) represents the upper level nesting, with the effects of lower level choices summarized in the effect of m_j (given that we do not observe any characteristics of lower level matches).

With homogeneous births, all covariates are tract characteristics and are not firm specific. In this case, if the expected number of births in a neighborhood, λ_j, takes the usual form

---

[11] That is, $C_j = \ln\left(Ce^{\eta_j}\right)$.

$\exp\left[\pi_j + (1-\sigma)\ln m_j\right]$, the problem may be equivalently estimated as a standard Poisson model of birth counts per tract (Guimarães, Figueiredo, and Woodward, 2000). We start by presenting ordinary Poisson results, but that leaves the problem that error draws for neighborhoods may be correlated with some neighborhood covariates. As such, we want instruments that influence these covariates but are exogenous to contemporaneous error draws. Assuming such instruments exist, we can estimate a (non-linear) moment condition. This also relaxes the Poisson assumption on mean and variance and takes care of possible over- or under-dispersion problems.

If we define

$$\upsilon_j \equiv B_j - \lambda_j = B_j - \exp\left[\pi_j + (1-\sigma)\ln m_j\right],$$

where B_j is the actual number of births and λ_j the expected number, the moment condition is

$$E\left[\upsilon_j \mid Z_j\right] = 0, \qquad (8)$$

where Z_j are instruments. We use non-linear generalized method of moments suggested in Windmeijer and Silva (1997) and Mullahy (1997) to estimate the parameters.

Data. We use Census Business Register [SSEL] which covers all US establishments. We analyze the location decisions of newborn SU's in Manhattan from 1992 to 1997. A birth is a new firm to Manhattan—a firm that did not exist in the 1992 SSEL but did in the 1997 one. We chose the SSEL in Census years because in those years additional information from the Census is available for firms; and more critically the records on existence and location are updated and most accurate for Economic Census years. We examine births in the 164 census tracts south of 90th Street, which cover 97.4% of all SU's. We exclude the 132 tracts north of 90th Street because we cannot accurately infer rent data per tract for them. Correspondingly, many of those tracts are not viable locations for advertising agencies.

Rents are based on zip code asking rents for class A office space in Manhattan in 1992; we utilize the rents in zip codes for which we have data and use a GIS rent-contour fitting routine to infer rents for all Census tracts in southern Manhattan (see Appendix A). Tract rental prices vary threefold within southern Manhattan.

Total births for SU's and MU's in Manhattan are given in Table 3. For our tracts south of 90th, the 1992 SU stock is 949 and the number of births is 502. A birth is a new company identifier, and as such includes buy-outs, so actual births may be for MU's.[12] We base 1992-1997 births on 1992 covariates. This specification implies that these births have naive expectations concerning economic magnitudes, which generally fits the data best. However, there is a lot of turnover in firms and their locations so firms born in, say, the last year (1997) may be basing their decisions on much more updated information than 1992. Thus based on 1997 Census data which contain a year of birth, we also look at a sample of SU births for the 148 establishments who in 1997 give their birth year as 1993 or 1994. [Yearly SSEL records for 1993 and 1994 show almost no births, reflecting the fact that birth records aren't updated until Census years.] We first show that results for the two samples are the same and then focus on 1992-97 births because the tract birth numbers are more numerous and hence less noisy.

For each tract, we construct measures of access of that tract to nearby advertising agencies, dividing the stock of agencies in 1992 relative to that tract into rings. We define 5 rings moving out in increments of 250 meters. Ring 1 is the count of existing SU advertising agencies in census tracts whose centroid is within 250 meters of the own census tract centroid. Mostly ring 1 is just the own census tract (but it can include up to 3 tracts). Ring 2 is the count of SU's in census tracts whose centroids are between 250 and 500 meters of the own tract, and so on for ring's 3-5, up to 1250 meters. Each additional ring typically contains 3-4 more tracts. We also experiment with other ring divisions and report one set of results.[13] Finally, we calculate the total number of commercial establishments for each tract and ring in Manhattan.

Instruments. In terms of identification, as noted earlier, by focusing just on Manhattan, we minimize the usual fixed effects problem in cross- MSA estimations, where localities have different unobserved characteristics in terms of labor markets, transport infrastructure, and regulatory-business-political climates. Now, the problem reduces to one of unobserved neighborhood characteristics within Manhattan, such as what the trendy eating and socializing spots for networking are, where construction is underway making face-to-face meetings more time costly, or where pick-pocketing, street people and the like are currently a more troublesome problem for employees and perspective clients. For these types of unobservable attributes, we

[12] There is more of a problem for MU's where there is a lot of corporate turnover. For example, for SU's about 10% of establishments have been around for 25 years or more. For MU's the number is much smaller, probably reflecting the general waves of buy-outs in the mid-1980's: the establishment may still be there but with a different owner, or firm identifier.

[13] Note southern Manhattan is about 2600 meters across and about 11,000 meters long.

believe it is realistic to assume they are constantly changing, so unobservable, relevant attributes in 1992 are uncorrelated with those 15 years later. As such, we choose historical variables from 1977 as instruments for 1992 covariates. We will report a variety of tests concerning this assumption below, but granting us that for now, the issue is what is endogenous and what are strong instruments.

In the empirical work, the coefficient subject to the greatest bias will turn out to be rents. The idea is that the effects of current neighborhood unobservables on businesses in general are capitalized into rents. Better conditions are reflected in higher rents, biasing the rent coefficient upwards. Given Manhattan's highly regulated housing market, commercial rent has strong instruments: total housing units in 1970, and share of those units in 1970 in buildings with 5 or fewer units. These variables reflect the potential, despite regulations, to convert residences to office spaces. In first stage regressions for rent, these variables have significant negative coefficients. We also control for distance to Rockefeller center as the commercial center of the city (which itself has no direct effects on advertising agency profits), but its effect in first stage regressions is weak.

The stocks of advertising agencies in rings are also potentially endogenous variables. Trendy local conditions not only influence rents and births today, but also may have enough persistence into the past to affect some of the accumulation of current stocks. The primary concern is with the own ring stock, since the error drawing is for the own tract in the own ring. Of course error drawings across tracts and rings may be correlated. While we will test for cross-tract spatial correlation and show that is not a compelling issue, we still instrument for all ring stock variables. For instruments we use 1977 values of ring stocks.[14] Use of these instruments raises two classical inter-related issues.

First, why could these be strong instruments, given we are asserting there is no persistence in unobservables influencing stocks in 1977 and 1992 and advertising agencies have shifted locations in Manhattan overtime, as documented earlier by comparing 1987 and 1997 location patterns? While there are distinct overall shifts, for many clusters there is also a reasonable degree of persistence, driven by moving costs. These include physical moving costs, loss of address recognition, and the costs of leaving (defecting from) an existing network (co-

[14] Actually we can only assign about 70% of establishments in 1977 to a 1992 census tract; and it is the counts of these 70% that we use.

ordination failure of moving clusters en masse). In general 1977 ring stocks are strong instruments, although surprisingly less so for the own ring.

Second, why are they valid instruments in identification? We note the results of many tests below, anticipating readers should a priori be suspicious. Besides formal and informal specification tests, these include (a) adding measures of other characteristics that may affect location choices and may persist over time, (b) looking at patterns of deaths to show birth patterns are not simply replicating death patterns, and (c) trying alternative instruments for the own ring stock.

For the last, an alternative instrument is based on a dramatic change in the way headquarters and advertising agencies interact. Not only have headquarters moved away from Manhattan over the last 20 years, but 20 years ago advertising agencies in Manhattan in their location choices clung to headquarter locations. In 1977 the simple correlation coefficient between numbers of advertising agencies and headquarters per tract was .82. By 1992, the corresponding correlation coefficient had fallen to .28 (recognizing patterns of residential versus commercial locations across Manhattan, generate positive correlation between any two business activities). With the internet, cheap air travel, cheap telephone service and the like, the advertising agency-headquarter interaction has changed. Now, day-to-day interactions are not face-to-face; and specifically in the highly intense design stage of advertising campaigns, agencies usually fly a team to headquarter locations away from Manhattan for some weeks. Given the historical need to be near headquarters and some persistence in locations of older advertising agencies, the correlation coefficient between 1992 SU location patterns and 1977 headquarter location patterns is .56, double the coefficient with contemporaneous headquarter location patterns. This makes 1977 headquarter counts by tract a potential ideal instrument for 1992 SU stocks, especially if one believes headquarters' location choices in 1977 were little influenced by where SU advertisers located, being more dependent on tract characteristics not relevant to advertising agencies (like access to financial services).

3. Results

In this section we start with results and identification issues for the basic birth specification in Table 4. Then we turn to Table 5 which pertains to identification as well. Table 6 explores effects of tract characteristics not included in the basic specification, and Table 7a and

7b examine robustness of the basic specification. Finally, we consider the problem of spatial correlation.

3.1 Basic SU Results

Table 4 contains our main results. The discussion is divided into three parts. First we discuss the effects of IV versus ordinary Poisson estimations, then we turn to the numerical interpretation of coefficients, and finally we discuss in great detail the validity of our instruments. The first two columns of Table 4 are for births in 1993 and 1994 (as recorded in the 1997 SSEL), which should be most sensitive to 1992 conditions; and the next two columns are for 1992-97 births. For 1993 and 1994 births, column 1 contains ordinary Poisson results (identical to ordinary conditional logit results); and column 2 contains the results of the non-linear generalized method of moments IV (GMM-IV) estimation. Comparing the two columns, the impact of IV estimation is concentrated on the own tract rent and the own (0-250 meters) SU ring stock coefficents, as expected. First, as expected, the absolute value of the IV coefficient on rent is much higher than the ordinary Poisson coefficient. Higher rents besides higher costs (negative effect) reflect better unobserved contemporaneous and expected future neighborhood amenities (positive effects) in the ordinary Poisson specification. Also as expected, the IV estimation of the own ring stock coefficient (.0218) is lower (by 35%) than the ordinary Poisson estimation. A larger number of agencies in the first ring may also reflect higher unobserved and somewhat persistent—from before 1992—neighborhood amenities affecting both new births and stocks in 1992.

In columns 3 and 4 for 1992-97 births, while the direction of bias in the rent coefficient is the same, there is no difference in the own ring stock coefficients between IV and ordinary Poisson estimation. We have three comments on this. First, additional births that occur from 1995 to 1997 are influenced by error drawings for those years which are now separated in time from the shocks that affected the accumulation of own ring stocks prior to 1992. Second, discussing bias in terms of absolute coefficients may be misleading. Given a discrete choice framework, coefficients aren't absolute profit function coefficients; below we show the degree of bias in terms of willingness to pay to have more agencies in ring 1 is larger for 1992-97 births than 1993-94 births: a 3.2 fold increase compared to a 2.4 increase. Third, the IV estimates of the own ring and rent coefficients are basically the same for 1992-97 births and 1993-94 births. In general we are going to rely on 1992-97 results because the sample of births is larger; but throughout results for the two samples are always very similar.

3.2 How to Interpret the Results

There are two interpretations of the coefficients of the estimated profit function, in the discrete choice framework. First, under a count model interpretation, coefficients idicate by what percent a change in a covariate raises or lowers the expected number of births in a tract.[15] To help evaluate relevant magnitudes, Table B-1 in Appendix B reports means and standard deviations of covariates. Second, we can monetize effects on profits utilizing the rent coefficient. We can calculate for each ring stock magnitude what percentage increase in rent per square foot of office space a firm would be willing to pay to have one more neighbor, or we can ask how many standard deviations increase in unit rent a firm would pay for a one standard deviation increase in neighbors. Later, for a very small sample of firms, we will actually look at the effects of covariates on *ex post* absolute profits, to give a sense of how the estimated coefficients of our discrete choice model compare in absolute magnitude to actual parameters.

Column 4 of Table 4 says that a 1% increase in rents would lead to a 2.6% decrease in expected numbers of births in a tract, indicating births within Manhattan are very sensitive to rents. In particular, a one standard deviation change (.17) in the rent variable leads to a 44% drop in expected births in a tract. This magnitude helps explain the shifting pattern of advertising agency locations in Manhattan in response to relative rent changes across tracts over the last 15 years, as noted earlier.

On the benefits of having more SU neighbors, column 4 shows that, as we move from the first ring out in 250 meter increments, the coefficients are .020, .023, .0042 and then zero beyond 750 meters. Ring 1 and 2 effects are always similar, but at 500 meters, the drop to the third ring is always large, with the coefficient being small and insignificant in IV estimation in column 4, although in column 2 it is significant at a 10% level. Any inferred networking effects end with ring 3. For the hypothesized networking effects, close spatial proximity is critical and thus so is having greater densities of establishments in commercial neighborhoods. Based on numbers of advertising agencies per ring in Appendix B and changes in ring areas as we move out, the

[15] Equivalently, in a multinomial logit framework, the effect is to change the probability of a birth by a particular percent (assuming the base probability of a birth in any tract is small)

number of advertising agencies per unit spatial area does not decline as we move to outer rings, so there is no sense that rings 4 and 5 have anything less to potentially offer the typical firm in the sample; they are just too far away. Interactions occur primarily within 500 meters, perhaps a 15-20 minute journey of elevator rides and walking during the day within Manhattan with its crowded conditions. This finding is robust to specifications of ring distance measures, as discussed below. And as noted earlier each additional ring encompasses 3-4 more tracts, which is a small increment.

One question concerns why the ring 1 effect doesn't dominate the ring 2? First, it could be that commuting costs are non-linear: people may be indifferent between a 5 minute and 12 minute walk, but not between those and a 20 minute one. Second, although we don't think there are shopping-center or market potential externalities, there may still be a poaching problem. While for a firm, it is great to have a network "member" 5-10 minutes walk down the block, having an agency right next door could "advertise" a neighbor's features to current clients who may choose to visit the agency occasionally.

In terms of magnitudes, if the number of neighbors in ring 1 increases by one, the expected number of births increases by 2%. For typical variations, the standard deviation of ring 1 neighbors is 10 and of ring 2 is 20 (and ring 3 is 49).[16] If we increase the number of immediate neighbors by one standard deviation that raises the expected number of births by 20% and if we increase the number in ring 2 by a standard deviation that raises the expected number of births in the own tract by 46%. Even for ring 3, with its point estimate of .004, the effect of a one standard deviation increase on expected number of births is 20%. These are very large absolute effects.

In terms of willingness to pay higher rent in order to enhance networking possibilities, column 4 indicates (the number given in square brackets) that a firm would be willing to pay 0.8 percent (.0198/2.57x100) higher rent per square foot to have one more neighbor in the first ring. That willingness to pay for a larger network is similar for ring 2. For ring 1, firms would be willing to pay just under a half a standard deviation increase in unit rent (or .078 of .17) for one standard deviation increase in immediate neighbors. Or if a firm moves from the fringe with no neighbors to the height of the action in our sample with 50 immediate neighbors, it would be willing to pay over 2 standard deviations in unit rental increases. Of course, that move is likely to

[16] Note the standard deviation and average of ring counts rise as we move out. It is expected since the ring area ($\pi(r_1 - r_2)(r_1 + r_2)$) rises as we move out where $(r_1 - r_2)$ is a constant (the 250 m. increment) but $(r_1 + r_2)$ rises with distance. Table B-1 in Appendix B reports the agency count and its standard deviation in our sample for the rings.

bring more neighbors in rings 2 and 3, as well. There is a substantial rent-networking trade-off in choosing locations for advertising agencies within Manhattan.

These networking-scale results are based on a linear formulation of ring counts. We experimented with a quadratic formulation and a log-linear one. We don't highlight these results because their IV versions are statistically weak.[17] But for ordinary Poissons both formulations "work". Under a log-linearity, the coefficients (and standard errors) for the first three rings are .391 (.066), .205 (.070), and .140 (.0603); these are very high scale elasticities. Quadratic scale effects are also large. For example for the first ring, the coefficients on the linear and squared terms are .0470 and -.000509 (both significant at 10% or better level). If we increase the number of neighbors by a standard deviation from 6 to 16, in the first ring expected births rise by 41 percent. Positive marginal effects persist until 45 neighbors, which interestingly coincides with the highest concentration of agencies in Manhattan (i.e., 51 agencies in the first ring).

Finally, the remaining variable is the number of commercial establishments of all types in a tract. This could represent some type of local agglomeration economy, or a market potential and shopping-center effect. But as we will show below, such effects do not extend beyond the first ring—a distance less than 250 meters. The variable could also represent the potential number of slots available for a new agency, where more slots offer better idiosyncratic matching of individual establishments to the neighborhood. in nested logit context, as discussed in subsection 2.2. Regardless, the effect is large: a 1 percent increase in the number of commercial establishments in the tract raises expected births by 0.5 percent. Under the latter interpretation, the coefficient implies a significant positive correlation between error drawings within tracts. We will later look at spatial correlation across tracts.

3.3 Identification

The instruments are listed in the footnote of Table 4. These instruments include 1970 housing market variables and counts of advertising agency and all commercial establishments for every ring in 1977. All 1992 covariates are treated as endogeneous, except for own ring total number of establishments. By all criteria total spaces for commercial establishments in a tract are exogenously given by New York zoning constraints and not sensitive to immediate market

[17] No coefficients are significant in the quadratic and only the coefficient for the first ring is significant in the log-linear formulation.

conditions. First stage regressions indicate the instruments are strong, partial and minimum *F* statistics are reported at the bottom of the table, and the partial R^2's average .50.

The chief issue is the validity of using 1977 ring stock variables as instruments, especially for the first ring. It is important to establish that despite some persistence in location patterns of advertising agencies over time due, for example, to moving costs which makes 1977 stocks possible instruments, tract unobservables change over time so 1992 residuals are uncorrelated with 1977 stocks. First, we note the simple correlation coefficient between 1992 and 1977 first ring stocks of SU's is only .60; the own ring stock first stage regression is the weakest with the lowest partial *F* (7.92); and the coefficient on the 1977 stock variable in the first stage regression for 1992 own tract ring stock only has a t-statistic of 1.44. Second, the Sargan tests cannot reject the validity of instruments; and overall Sargen statistics in columns (2) and (4) of the table are very low, indicating that the specification and our assumption that current errors are orthogonal to past stocks are valid. Nevertheless, we conducted many other experiments to test validity of instruments.

First as an informal test we also 1977 stocks for five rings to the ordinary Poisson in column 3 of Table 4. The results are in the text table where, the first ring reports ring stock coefficients for the ordinary Poisson and the second those coefficients when 1977 ring stock instruments are added in. If 1977 stocks are not valid instruments (are correlated with error

	'92 Ring 1	**'92 Ring 2**	**'92 Ring 3**	**'92 Ring 4**	**'92 Ring 5**	**'77 Ring 1**	**'77 Ring 2**	**'77 Ring 3**	**'77 Ring 4**	**'77 Ring 5**
Base case (Table 4)	.021** (.0043)	.017** (.0026)	.0039** (.0014)	-.0015 (.0016)	-.0022 (.0015)					
Instruments added	.023** (.0049)	.017** (.0030)	.0052** (.0018)	.0034 (.0023)	-.0015 (.0019)	.00046 (.0034)	.0035 (.0028)	-.0031* (.0017)	-.0037** (.0018)	-.0025 (.0016)

terms) and 1992 stocks are also correlated with error terms, the coefficients of 1992 stocks should decline. The table shows for the key variables for the first three rings, no decline occurs. Moreover, the 1977 stocks have zero coefficients for the first two rings. We also divided the 1992 stocks into new (post 1982 birth date) versus old firms, inserting each as separate sets of covariates. In ordinary Poissons, the former strongly dominate the latter because the new stocks are more likely to be correlated with current shocks. In IV estimation, with so many variables the results are not statistically strong, but the pattern of dominance by new stocks disappears as it should, because IV estimation removes any correlation of new stocks with current shocks.

In Table 5, we explore the validity of instruments and our approach further. Column 1 repeats the base case from Table 4, column 4. In column 2 of Table 5, we replace the 1977 own ring stock instrument with the own ring count of headquarters in 1977 as discussed earlier. That leads to deterioration in the Sargan test statistic and an even weaker first stage *F*-statistic (6.1) for the own ring stock, although the Sargan p-value remains a satisfactory .41. The rent and relevant ring stock coefficients are both somewhat lower under this specification with for example the own ring stock coefficient dropping 18%; however the willingness to pay magnitudes for additional neighbors are similar to the base case (even higher for the own ring).

Another way to view the concern with using the 1977 own ring stock variable as an instrument is that births in our context may simply replace deaths --old agencies-- in a tract, so location patterns simply replicate over time. Higher stocks mean higher deaths and hence almost mechanically higher births. We have already shown in fact that SU location patterns changed significantly between 1977 and 1992. However we can tackle the issue directly by examining deaths. We estimate an equation for agency deaths from 1992-1997 using the same set of covariates. The results are reported in columns 3 and 4 of Table 5. Deaths are agencies that disappear entirely (not movers within Manhattan). Based on Davis, Haltiwanger, and Schuh (1996), individual deaths occur for idiosyncratic reasons, in particular special circumstances involving the entrepreneur. As such, agencies' deaths by tract should be roughly proportional to tract stocks, so the first ring coefficient will be significant. [We note the coefficient of 0.076 on the first ring stock suggests that if we go from zero to mean stock per tract of about 6, deaths increase in the tract by 45%. That means about half of the tract stock dies, which is what happens on average.] However, if births simply replace deaths so tract patterns are replicated over time, the death equation estimates should mimic the birth ones. The ordinary Poisson (column 3) and in particular the IV estimates (column 4) show this is definitely not the case. In column 4, the effects of rent, total establishment, ring 2, and ring 3 disappear completely, neither mimicking effects for births nor offering instructive results on deaths. The death equation fails the Sargan specification test, which as a joint test on instruments and specification, suggests the death model is not well specified.

3.4. Robustness to inclusion of related activities.

Following the idea of backward and forward linkages, in Table 6 we explore the effects of proximity in Manhattan to activities other than advertising agencies. Column 1 repeats our base case estimates from column 4 of Table 4. Column 2 shows the effect of dropping the control for total commercial activity in the own tract. Doing so strongly increases the coefficient of and willingness-to-pay for the first ring stock of SU agencies. This suggests that the control does capture important aspects of having more commercial activities in the own tract, whether they reflect some type of local agglomeration economy beyond own industry, or greater choice among own tract commercial rental properties in the nested logit approach discussed in section 2.2. Column 3 shows that having more commercial establishments in rings beyond the first, or own tracts, generates no benefits. This suggests total commercials establishments may not reflect information spillovers or market potential externalities since it seems unlikely such benefits dissipate so quickly. Instead they may reflect an underlying nested logit choice context.

In the remaining columns we explore more specific aspects of access to other commercial enterprises. Initially we looked at establishments of multi-unit advertising agency firms. Ordinary Poisson results suggest that having MU's nearby help in the first ring (significant coefficient of .027) and hurt in the second; but IV results give small and insignificant coefficients (first ring coefficient of .0045). This fits with results in the Appendix that show multi-unit agencies do not derive benefits from SU's.

We then turned to the key issue. For advertising agencies, to what other commercial activities is access important, where our ring stock variables might be capturing these effects rather than information spillovers? We had difficulty finding significant effects and widened the net and strengthened any effects by combining the first two rings and defining possible effects for these other activities to occur in 0-500 meters. In the discussion of instruments, we already explained that by 1992 access to buyers, or headquarters, is unrelated to location choices within Manhattan. Thus the coefficient (not reported in Table 6) on the count of headquarters within 500 meters is negative and insignificant in both ordinary Poisson and IV estimation.
In column 4 of Table 6, we examine access to broadcasters who place advertisements. The effect of broadcasters is the only significant one in these experiments. We measure broadcasting scale by total employment in broadcasting; this works better than using a count of etablishments given the vast dispersion in broadcast establishment sizes. The IV results show that a one standard

deviation increase in broadcasting employment (1785) raises expected births by 18%, a noticeable effect (although the ratio of standard deviation to mean is high (3.8) reflecting zeros in many rings, with very high concentrations in a few). Inclusion of this variable leaves other coefficients unaffected. Finally we turn to a key input to advertising agencies: graphic services. In IV estimation, the partial *F* in first stage regressions for this variable (after adding 1977 stocks on headquarters, broadcasters and graphics design to the list of instruments) is a dismal 2.26. Thus, we report only ordinary Poisson results, which suggest no effect of having more graphic services nearby. In summary, we find no evidence that inclusion of other suspected relevant local activities diminishes the estimated benefits of having more advertising agencies in successive rings.

3.5 Robustness to specification

In Table 7, we carry out other robustness checks for our basic specification and sample of tracts. First, in Table 7a, we restrict the sample of tracts to 138 tracts that have ever had a presence of advertising agencies (SU or MU) since 1977. That eliminates 26 tracts that may not be relevant to advertising agency location decisions. Coefficients in column 1 are similar to those in Table 4. While absolute values tend to fall (the rent coefficient is about 25% lower and the ring 1 stock variable coefficient is about 11% lower) as expected[18], the willingness to pay for neighbors rises modestly.

In columns 2 and 3 of Table 7, we consider the absolute magnitudes of our estimated, normalized coefficients. For the 30 SU advertising agencies in Manhattan for which 1992 Business Expenditure Survey data exist, we calculate profits as sales minus all operating expenses including rent. We regress the log of profits on the same covariates as our basic births model in order to estimate non-normalized parameters of the profit function directly. Given the small sample size, we focus on OLS results. While under a nested logit specification to the location-count problem, we had to control for the total number of establishments in the tract, that argument is no longer relevant per se for this direct estimation of profits. However, the number of establishments does control for possible general local agglomeration economies, as noted earlier. In columns 2 and 3 of Table 7a we show OLS results with and without the control; the coefficient on this control is weak (relatively speaking in a context where all results are weak

[18] In general, the coefficients in discrete choice models— the generalized exponential model in our case—are scaled by the standard deviation of the residuals. As we excluded the tracts with no births, we expect that the standard deviation of the residuals and therefore the magnitude of coefficients decrease.

with the small sample size). It is not a compelling control in a small sample context, where we want as few parameters as possible. In columns 2 and 3, point estimates of coefficients for total profits are larger in absolute terms than the count results. But, the pattern of coefficients is similar to those in our basic birth model, with a large negative rent coefficient and high benefits of near neighbors, which decline sharply with distance. The results hint at what absolute coefficients might be, compared to our count estimations. In the OLS profit function in column 3, coefficients are at least twice ordinary Poisson coefficients for rents and the first two ring variables. Column 4 shows 2SLS results where we instrument for rents and the first 2 ring stock variables.

In column 5 of Table 7a, we turn our specification of ring distances. We defined rings by equal increments of distance to reflect equal changes in contact costs. Thus, the contact (commuting) costs to an agency in the second ring are twice those for the first ring. We believe this in the best way to specify distance decay. However one could define, for example, rings to have equal area, rather than equal incremental distances from the own tract. In column 2, we define equal area rings of 0-350, 350-500, 500-610 and 610-707 meters. Results do differ somewhat from the equal distance specifications. First, the ring effect in moving from ring 1 (ending at 350 meters) to ring 2 increases significantly. This suggests the poaching issue may be stronger than hypothesized above. While the scale effects persist for 500-610 meters, they disappear at 610 meters, shorter than our 750 meters before.

The final issue is that we have measured advertising agency networking and information spillover effects based on just a count of enterprises. The idea is that each enterprise represents a source of information spillovers and a networking opportunity. It could be that employment counts are a better representation of networking opportunities, but how can we distinguish the two? One way is to decompose total SU ring employment into a count of establishments and average employment size to see if they yield similar effects. Such a decomposition serves double duty, allowing for the possibility that, even if establishment counts accurately record the opportunities for networking, quality of neighborhood establishments may matter. If size is a measure of quality as discussed in section 2.1, this decomposition allows us to both control for sources of externalities and quality of neighbors.

The basic problem in proceeding with a decomposition is that for IV estimation, while we can still instrument for establishment counts, we have no strong instruments for average establishment size. Adding to our instrument list 1977 variables on average SU size by ring produces first stage regressions for average establishment size with F-statistics under 2.0. Corresponding if we make the spillover covariates total employment counts in 1992, in first stage regressions all variation in total employment counts is explained just by 1977 establishment count variation. We report on two sets of results.

First in Table 7b, column 1 for IV estimation we replace establishment counts with employment counts of advertising agencies (log employment) in rings. Given our comments on first stage regressions in the previous paragraph, the partial *F*-statistic for the own ring employment stock regression is only 3.7. This makes it hard to place much weight on the results. The results show positive employment effects for the first three rings (now in elasticity form: percent changes in births for a percent change in SU employment stocks).

Second, we report on two decomposition attempts for total employment, for ordinary Poissons only, since we have no instruments for establishment size measures. While the basic decomposition is into the count of agencies and average employment per agency to distinguish the effects of more firms versus either more employees and/or larger and higher quality firms, averages can be a noisy measure of quality driven by outliers. Thus we also experimented with median firm employment by ring, as the quality-establishment size measure. For the ordinary Poisson estimation, we define just three rings (without employment size measures, coefficients on SU stocks for rings 1, 2 and 3 are respectively .0214, .0149, and .00265, similar to results when five rings are specified). Adding median SU employment by ring in column 3 of Table 7b shows no effect on establishment count coefficients and yields insignificant coefficients for the employment size measures. However, using average size measures reduces the employment count coefficients for rings 1 and 2 by 18 and 30% in column 4; but we note average size and establishment counts are positively correlated in the data (simple correlation coefficients of about .3-.4). Average employment size variables are significant for ring 1 and almost so for ring 2. So there is at least a hint that employment sizes as well as establishment counts may matter. We can't sort this out for IV estimation.

Finally in Appendix B, we present results for MU's where the stock is 132 establishments and the number of births is 67. Table B-2 shows that estimated coefficients for MU's have similar patterns of signs and significance to those of SU's. The discussion in the Appendix indicates that, just as SU's benefit from nearby SU's and not MU's, MU's benefit from nearby

MU's and not SU's. This separation implies that MU's may for example interact with other MU's to share information about media purchases, of which SU's do relatively less; while SU's may interact with each other at the creative design stage across firms. For that, MU's may rely on intra-firm interactions at the design stage.

3.5 Spatial Correlation.

When using data at this fine level of detail, and in addition when asserting within tract correlation of error terms, the issue of spatial correlation of error terms across tracts arises. Before addressing that directly, note that in Table 4, when we instrument for ring variables, we instrument for them all. So if the shock to the own ring affects ring stocks and shocks are correlated across tracts, then the shock to the own ring could be correlated with ring 2 or ring 3 covariates. Instrumenting with historical variables takes care of that problem; even if the instruments are historically spatially correlated, the concern is that the shocks today are independent of these historical variables.

The key issue with spatial correlation is the potential for biased estimates of standards errors. Thus we checked for spatial correlation, using a Moran I test on the residuals from both the ordinary Poisson and the GMM models in Table 4, columns (3) and (4). The Moran I test examines whether a regression of neighbors' average residuals on own tract residuals is significant. Near neighbors in the averaging are given a weight 1 and others a weight 0. In denoting what is a near neighbor, the usual procedure is to think of each tract as having 8 neighbors on a lattice. We use a radius from the own tract that gives an average of 8.3 neighbor tracts as neighbors. For this radius, the hypothesis of no correlation cannot be rejected at a p-value of .10 for an ordinary Poisson. The result for GMM coefficients is similar—no rejection with a p-value of .11.[19] Thus within our data, while, given tracts are neighborhoods, there may be within tract correlation, the problem of cross tract correlation does not seem compelling.

[19] Only if we drop the radius so the average number of neighbors is 3.3, can the hypothesis of no correlation be rejected (p-value of .015). We also regressed residuals of pair-wise neighbors on each other, getting exactly the same pattern of results. For those we also looked at a sample requiring all pairs to both have counts greater than 1 (worrying about the discreteness problem with count data). The same results apply.

4. Conclusions and Extensions

The fundamental findings in the paper are (1) at a micro spatial level, scale externalities are very large but they dissipate very quickly with distance for advertisers and are gone by 750 meters and (2) benefits of being in a better neighborhood are capitalized into land rents. The effects seem to strongly suggest networking, or information “spillover” effects. The results raise two other key points. First use of wage equations to estimate urban agglomeration economies ignores the fact that some portion of benefits are capitalized into both the overall level of commercial rents in a city and rent variation across space within commercial sectors. Second the finding of quick spatial decay of information spillover benefits for advertising agencies begs the question of what one measures with more aggregate data, such as MSA scale effects. Is it labor market externalities (operating at the level of the entire labor market) or is it that having more advertising agencies in a city allows a greater choice of clusters within a city?

The idea that firms within a city in choosing locations are trading-off rent costs versus the benefits of being “at the center of action” in a large cluster versus operating on the fringes raises the issue of heterogeneity. Which firms are choosing to pay higher versus lower rents in order to have better neighborhood networking – are these higher or lower quality firms? In an earlier version of the paper, we looked at this question in detail, but with a very limited sample. As such results are suggestive results, but there is limited precision, so our presentation is brief.

In introducing observed heterogeneity, we assume from the model that better quality firms are larger, so we can treat size as an index of quality. We can’t introduce heterogeneity for births, because we have no ex ante information on their quality (and the firms themselves may not have learned their quality); and ex post size measures are likely to be influenced by contemporaneous error drawings. However we have a sample of 95 movers who change location between 1992 and 1997 within Manhattan. These experienced firms have informed notions of their own quality; and, for 82 of them, we have a size measure of quality-- payroll in their prior location in 1992. That variable will be uncorrelated with the error drawings in the new location. We note payroll and employment information from social security records are the only non-imputed census variables on firm economic magnitudes available for more than a handful of firms. We use payroll (where there is a uniform wage for uniform workers in Manhattan), to allow better quality firms to also hire better quality, or higher paid employees, but there is little difference in results.

To incorporate firm heterogeneity into the empirical estimation, we follow the approach in Berry (1994) and Bayer, McMillan, and Rueben (2004) and postulate a heterogeneous coefficients model. In the profit function appearing in (6) we assume the coefficient vector ϕ for neighborhood covariates now takes the form

$$\phi^k = \phi_0 + \phi_1 z^k . \qquad (9)$$

ϕ^k are firm k 's specific coefficients on the X_j covariates, where those coefficients vary with firm characteristics, z^k , which in this case will be a measure of firm quality. For scale effects in equation (6), we are allowing $\ln E[\cdot]$ to change as either γ or the expected *u* drawing from equation (5) vary with firm quality; for rental effects we are allowing α to vary. Once we introduce quality issues, in our framework we should also ask whether firms value quality of neighbors, as well as counts. For an incoming firm, payroll levels of existing firms are a signal as to neighborhood quality. As such, we use the median firm payroll by ring for SU's in 1992 as a measure of neighborhood quality, recognizing this is an approximation.[20] As we already know, however, for ordinary Poisson's only the first ring has significant median neighborhood size effects and in IV work we have weak instruments. In what we discuss, we only control for ring one quality of neighbors.

Given neighborhood characteristics are endogenous and we have firm specific characteristics, we employ a two stage identification strategy. First, we estimate (7) with tract fixed effects, so the basic logit equation for firm *k* considering location *j* is

$$P_j^k = \frac{\exp[d_j + \phi_1 z^k X_j]}{\sum_i \exp[d_i + \phi_1 z^k X_i]} \qquad (10)$$

In (10), d_j is a neighborhood dummy variable, which controls for all neighborhood characteristics, observed or not. As long as the z^k are exogenous, the ϕ_1 are then identified. But that leaves the base coefficients, ϕ_0 , for the X_j covariates. These are given by the equation

$$d_j = b_j + \phi_0 X_j , \qquad (11)$$

[20] The issue is characterizing quality by a single measure of size. For example, this is a correct measure if the stock of firms is spread across neighborhoods so as to perfectly segment the line representing the quality types in the market. With perfect segmentation, the lowest quality firm in a neighborhood is at least the same quality as the highest quality firm in the next best quality neighborhood. Then greater size signals a first degree stochastic dominant improvement in the distribution.

where the LHS dummy variables are estimated in (10). To identify the ϕ_0, given the X_j are correlated with unobserved neighborhoods contemporaneous attributes, equation (11) is estimated by 2SLS, using as instruments the variables we used before in estimating the moment condition (8).

In estimation, because of the limited sample, to sharpen the reported results, we truncate ring effects at 3 rings. Results are given in Table 8. The first two columns give ordinary logit results for movers. Column (1) presents results without heterogeneity, where scale is for 3 rings and quality for 1. Column (2) presents the ordinary conditional logit results for the full model in (7), amended by (9) to allow for heterogeneous coefficients, so all variables in column (1) are now additionally interacted with the quality measure— payroll size in the prior location. If we evaluate these at median size of movers, results in column (2) are similar to those in column (1) on all variables. For that comparison,[21] we use a "synthetic" median, which is 4.75; this is the log of the mean size of the 10 firms around the median. However there are enormous differences by firm quality. In particular ordinary logit results suggest better quality firms are much less sensitive to rent differentials and much more sensitive to immediate ring 1 neighbors. We explore this next in columns (3) and (4), where issues of endogeneity are handled.

In column (3) we present results of the tract fixed effect logit model in equation (10). Notice the coefficients for these interactive variables are similar to those in column (2), but with stronger rent and ring 1 scale coefficients. In column (4), we recover the ϕ_0 coefficients by estimating equation (11) by 2SLS. For this, the sample size is only 40 tracts which raises a small sample problem, so in column (5) we also give the OLS results. To evaluate effects of any variable, we must combine the results in columns (3) and (4).

Results indicate that higher quality firms are less concerned about rent costs. The strongly negative rent coefficient for a small firm declines in absolute magnitude to -.36 for the median size mover. This could indicate that there is also a fixed cost portion to rent which effect diminishes with firm size. Second, higher quality firms benefit more from scale externalities, at least in ring 1. The ring 1 scale effect becomes positive at a very small firm and then rises sharply, so that at median size it is a high .034 (cf. Table 4 coefficient of .02) and at median size

[21]At median size, the establishment, rent, and ring 1-3 scale and ring quality (heterogeneous) coefficients are respectively .21, -.36, .022, .015 and .0075, while those in column (1) are .30, -.093, .020, .018, and .0059, indicating here a formulation without heterogeneity tends to capture the coefficients of the "representative" firm .

plus one standard deviation of size it is a huge .078. High quality firms are much more likely to choose neighborhoods with more SU's. Given the rent effect declines with firm quality, the willingness to pay in terms of rents for more neighbors is very high for higher quality firms. Point estimates suggest the median firm is willing to pay a 9.4% increase in unit rent to have one more neighbor, but with our limited sample size we lack precision. Finally, in the Table, results in all columns on neighborhood quality for our movers, as measured by the median payroll size of existing establishments in 1992 in the first ring, suggest quality doesn't matter. Of course we are looking within New York, which nationally has a high end cut of larger and presumably higher quality firms (Arzaghi, 2005).

References

Arzaghi, Mohammad (2004), “A Model of Knowledge Sharing and Network Formation Among Advertising Agencies,” Working Paper, Brown University.

Arzaghi, Mohammad (2005), “Quality Sorting and Networking: Evidence from the Advertising Agency Indsutry,” Working Paper, Brown University.

Bala, Venkatesh; Goyal, Sanjeev (2000), “A Noncooperative Model of Network Formation,” *Econometrica*, 68, 1181-1229.

Bayer, Patrick; McMillan, Robert; Rueben, Kim (2004), "Residential Segregation in Equilibrium," Yale University mimeo.

Berndt, Ernest R; Alvin J. Silk (1994), “Costs, Institutional Mobility Barriers, and Market Structure: Advertising Firms as Multiproduct Firms,” *Journal of Economics and Management Strategy*, 3, 437-450.

Berry, Steven (1994), "Estimating Discrete-Choice Models of Product Differentiation," *Rand Journal of Economics*, 25, 242-262.

Davis, S., J. Haltiwanger, and S. Schuh (1996), *Job Creation and Destruction*, Cambridge, MIT Press.

Dudey, Marc (1990), “Competition by Choice: The Effect of Consumer Search on Firm Location Decisions,” American Economic Review, 80(5), 1092-1104.

Duranton, Gilles; Overman, Henry (2004), "Testing for Localization Using Micro Economic Data," LSE Mimeo.

Ellison, Gllen; Glaeser, Edward (1997), "The Geographic Concentration of Industry: A Dartboard Approach," *Journal of Political Economy*, 105, 889-927.

Glaeser, Edward; Gyourko, Joseph (2002), "The Impact of Zoning on Housing Affordability," Harvard Institute of Economic Research Paper # 1948.

Goyal, Sanjeev; Moraga-Gonzalez, Jose Luis (2001), “R&D networks,” *RAND Journal of Economics*, 32(4), p 686-707.

Guimarães, Paulo; Figueirdo, Octávio; Woodward, Douglas (2000), "A Tractable Approach to the Firm Location Decision", *Review of Economics and Statistics*, 85, 201-204.

Hameroff, Eugene (1998), *The Advertising Agency Business*, New York: McGraw-Hill.

Henderson, J. Vernon (2003), "Marshall's Scale Externalities," *Journal of Urban Economics*, 53(1), 1-28.

Holmes, Thomas (2005), “The Location of Sales Offices and the Attraction of Cities,” *Journal of Political Economy*.

Jackson, Matthew; Wolinsky, Asher (1996), “A Strategic Model of Social and Economic Networks,” *Journal of Economic Theory*, 71, 44-74.

Jaffe, Adam B.; Trajtenberg, Manuel; Henderson, Rebecca (1993), "Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations", *Quarterly Journal of Economics*, 108, 577-598.

Lucas, Robert E. Jr. and Estaban Rossi-Hansberg (2002), “On the Internal Structure of Cities,” *Econometrica*, 70, 1445-1476.

McMillen, Daniel P. and Larry D. Singell, Jr. (1992), “Work Location, Residence Location, and The Intraurban Wage Gradient,” *Journal of Urban Economics*, 32, 195-213.

Mullahy, John (1997), "Instrumental Variables Estimation of Count Data Models, With an Application to Models of Cigarette Smoking Behavior," Review of Economics and Statistics, 79(4), 586-593.

Rosenthal, Stuart; Strange, William (2004), "Evidence on the Nature and Sources of Agglomeration Economies,” in *Handbook of Regional and Urban Economics*, J.V. Henderson and J-F Thisse (eds.), North Holland.

Rosenthal, Stuart; Strange, William (2003), "Geography, Industrial Organization and Agglomeration," *Review of Economics and Statistics*, 85, 178-188.

Schemetter, Bob (2003), *Leap! A Revolution in Creative Business Strategy*, Hoboken, NJ: John Wiley & Sons, Inc.

Windmeijer, Frank; Santos Silva, Joya (1997), "Endogeneity in Court Data Models: An Application to Health Care," *Journal of Applied Econometrics*, 12, 281-294.

Table 1. 1992 Economic Census Numbers on Concentration in Advertising*

	Advertising agencies		All other business services	
	Share of total	**No. of units**	**Share of total**	**No. of units**
Counts of establishments with under 5 employees	58%	6890	52%	133,380
Sales by establishments with under 5 employees	14%	6890	10%	133,380
Counts of establishments with 100+ employees	1.3%	156	4.1%	10,590
Sales by establishments with 100+ employees	33%	156	40%	10,590
Counts of single-unit firms	90%	12,453	81%	236,212
Sales by Single-unit firms	55%	12,453	45%	236,212
Sales by multi-unit firms	45%	581 firms 1481 est.	55%	12,146 firms 56,460 est.
Sales by multi-unit firms with 10+ establishments	22%	17 firms 408 est.	33%	771 firms 31,325 est.

* The establishment numbers cover just "operating establishment". The firm numbers cover all establishments. Numbers are from the 1992 Census of Services.

Table 2. New York City's Role in the Advertising agency Industry 1992

Share of US establishments	7.3%
Share of US employment	20%
Share of US receipts	24%
Share of media billing	31%

Table 3. New York City Single Unit versus Multi-Unit Establishments

	1997 Stock	**Born Before**		**'92-97 Births**
		'73	**'88**	
single unit est.	1088	104	357	545
multi-unit est.	149	3	77	67

Table 4. Birth Models for SU's.

	Birth '93-'94 Poisson	Birth '93-'94 IV-GMM	Birth '92-'97 Poisson	Birth '92-'97 IV-GMM
Ln(total no. of	.411**	.530**	.462**	.498**
establishments)	(.137)	(.216)	(.0756)	(.117)
Ln(rent/sq. ft.)	-1.72**	-2.74*	-.819**	-2.57**
	(.572)	(1.46)	(.289)	(.773)
Stock SU's	.0334**	.0209*	.0206**	.0198**
0-250m.	(.00817)	(.0119)	(.00430)	(.00758)
[Will-to-pay/sq. ft. (%)]	[1.9]	[.76]	[2.5]	[.77]
Stock SU's	.0147**	.0191**	.0167**	.0228**
250-500m.	(.00494)	(.00953)	(.00264)	(.00538)
Stock SU's	.00451*	.00648	.00387**	.00419
500-750m.	(.00250)	(.00500)	(.00137)	(.00259)
Stock SU's	.000702	-.00336	-.00149	-.00149
750-1000m.	(.00281)	(.00639)	(.00158)	(.00274)
Stock SU's	-.00577**	-.00519	-.00219	-.00241
1000-1250m.	(.00272)	(.00605)	(.00147)	(.00251)
N	164	164	164	164
Pseudo R^2	.343	—	.504	—
Sargan [p-value]	—	3.88 [.694]	—	1.71 [.945]
1st stage avg. R^2	—	.50	—	.50
avg. F; (min F)	—	13.0; (7.9)	—	13.0; (7.9)

The instrument list is distance to Rockefeller Center, 1970 log of total housing units, 1970 share of housing units in buildings with less than 5 units, stocks of SU advertising agencies in each of the 5 rings in 1977, total count of all establishments in each of the first four rings in 1977, and 1992 total number of establishments in the own tract.

Table 5. Alternative Instruments and Births versus Deaths

	Births '92-'97 IV-GMM	Births '92-'97 CA Instrument for the First Ring IV-GMM	Deaths '92-'97 Poisson	Deaths '92-'97 IV-GMM
Ln(total no. of	.498**	.530**	.366**	.0600
establishments)	(.117)	(.110)	(.101)	(.209)
Ln(rent/sq. ft.)	-2.57**	-1.98**	.164	-.0121
	(.773)	(.578)	(.364)	(.819)
Stock SU's	.0198**	.0167**	.0598**	.0756**
0-250m.	(.00758)	(.00626)	(.00578)	(.0140)
[Will-to-pay/sq. ft. %]	[.77]	[.84]		
Stock SU's	.0228**	.0169**	.00822**	.00864
250-500m.	(.00538)	(.00304)	(.00335)	(.00617)
Stock SU's	.00419	.00401**	-.000208	-.00182
500-750m.	(.00259)	(.00147)	(.00176)	(.00273)
Stock SU's	-.00149	-.00212	.00447**	.0111**
750-1000m.	(.00274)	(.00181)	(.00188)	(.00352)
Stock SU's	-.00241	-.000949	-.00373**	-.00182
1000-1250m.	(.00251)	(.00167)	(.00183)	(.00199)
N	164	164	164	164
Pseudo R^2	—	—	.549	—
Sargan [p-value]	1.71 [.945]	6.11 [.411]	—	13.0 [.042]
1st stage avg. R^2	.50	.36	—	.50
avg. F; (min F)	13.0; (7.9)	11.3; (6.1)	—	13.0; (7.9)

Except for column 2, the instrument list is distance to Rockefeller Center, 1970 log of total housing units, 1970 share of housing units in buildings with less than 5 units, stocks of SU advertising agencies in each of the 5 rings in 1977, total count of all establishments in each of the first four rings in 1977, and 1992 total number of establishments in the own tract. In column 2, the office rents and first ring counts are instrumented using the distance to Rockefeller Center, two 1970 housing variables, and 1977 first ring headquarter counts. The other rings counts are assumed exogenous.

Table 6. Other Covariates

	Base Case IV-GMM	No Control for Total Estab. IV-GMM	Total Estab. in Other Rings IV-GMM	Broadcasting Employ. IV-GMM	Graphic Design Poisson
Ln(total no. of	.498**	—	—	.563**	.463**
Establishments)	(.117)			(.120)	(.0756)
Ln(rent/sq. ft.)	-2.57**	-3.60**	-2.34**	-2.60**	-.701
	(.773)	(1.07)	(1.22)	(.850)	(.331)
Stock SU's	.0198**	.0321**	.0214**	.0221**	.0191**
0-250m.	(.00758)	(.00943)	(1.22)	(.00916)	(.00476)
[Will-to-pay/sq. ft. %]	[.77]	[.99]	[.91]	[.85]	[2.7]
Stock SU's	.0228**	.0341**	.0213**	.0190**	.0151**
250-500m.	(.00538)	(.00631)	(.00738)	(.00530)	(.00339)
Stock SU's	.00419	.00874**	.000769	.00281	.00358**
500-750m.	(.00259)	(.00322)	(.00604)	(.00261)	(.00142)
Stock SU's	-.00149	.00450	—	-.00298	-.00162
750-1000m.	(.00274)	(.00416)		(.00292)	(.00160)
Stock SU's	-.00241	-.00484	—	-.000954	-.00192
1000-1250m	(.00251)	(.00331)		(.00290)	(.00151)
Ln (total est)			.379**		
0-250m			(.178)		
Ln (total est.)			-.0529		
250-500m			(.0845)		
Ln (total est.)			.251		
500-750m			(.370)		
Broadcast employ.				.000103**	
0-500m				(.000035)	
Graphic services					.00152
0-500m					(.00207)
N	164	164	164	164	164
Pseudo R^2	—	—	—	—	.505
Sargan [p-value]	1.71 [.945]	6.03 [.537]	3.29 [.655]	1.57 [.954]	—
1st stage avg R^2	.50	.60	.33	.55	—
avg. F; (min F)	13.0; (7.9)	20.0; (13.2)	8.3; (6.1)	16.3; (7.4)	—

For column 3, the instrument list is distance to Rockefeller Center, 1970 log of total housing units, 1970 share of housing units in buildings with less than 5 units, and stocks of SU advertising agencies in each of the first three rings in 1977. The log of total count of all establishments in each of the first three rings in 1992 are considered to be exogenous.

Table 7a. Robustness: Sample, Coefficient Magnitude, and Rings Definition

	“relevant” tracts IV-GMM	Profits ‘92 OLS	Profits ’92 OLS	Profits ’92 2SLS	Equal area rings IV-GMM	
Ln(total no. of	.458**	.464	—	—		.569**
Establishments)	(.110)	(1.01)				(.0934)
Ln(rent/sq. ft.)	-1.96**	-2.94	-2.30	-2.80		-1.90**
	(.719)	(2.78)	(2.37)	(3.03)		(.823)
Stock SU’s	.0176**	.0743	.0905**	.1088*	**0-350m**	.0129**
0-250m.	(.00747)	(.0561)	(.0433)	(.0580)		(.00521)
[Will-to-pay/sq. ft. %]	[.90]					[.68]
Stock SU’s	.0178**	.0331	.0322	.0836**	**350-500m**	.0209**
250-500m.	(.00524)	(.0255)	(.0250)	(.0354)		(.00510)
Stock SU’s	.00399*	-.0248**	-.0249**	-.0325**	**500-610m**	.0107**
500-750m.	(.00224)	(.0110	(.0108)	(.0133)		(.00496)
Stock SU’s	-.00270		—	—	**610-707m**	-.00159
750-1000m.	(.00254)					(.00465)
Stock SU’s	-.00146		—	—		
1000-1250m.	(.00224)					
N	138	30	30	30		164
Adjusted/Centered R^2	—	.122	.150	.132		—
Sargan [p-value]	1.84 [.934]		—	—		5.29 [.508]
1st stage avg. R^2	.48		—	.55		.50
avg. F; (min F)	10.1; (5.4)		—	4.5; (3.5)		14.3; (9.6)

The instrument list for column 1 is distance to Rockefeller Center, 1970 log of total housing units, 1970 share of housing units in buildings with less than 5 units, stocks of SU advertising agencies in each of the five rings in 1977, total count of all establishments in each of the first four rings in 1977, and 1992 total number of establishments in the own tract. For column 4, ring sizes are adjusted in calculating instruments; we also only use stocks of SU’s for four rings.

Table 7b. Robustness: Scale Measures

	Births 92-97 IV-GMM		Births 92-97 Median emp. Poisson	Births 92-97 Average emp. Poisson
Ln(total no. of	.248**		.378**	.339**
Establishments)	(.0987)		(.0746)	(.0740)
Ln(rent/sq. ft.)	-1.07**		-.801**	-.912**
	(.432)		(.288)	(.280)
Ln(tot. SU emp.)	.238**	**Ln(SU emp.)**	.0659	.153**
0-250m.	(.104)	**0-250m.**	(.00423)	(.0570)
Ln(tot. SU emp.)	.352*	**Ln(SU emp.)**	.0292	.112*
250-500m.	(.215)	**250-500m.**	(.0881)	(.0632)
Ln(tot. SU emp.)	.190*	**Ln(SU emp.)**	.148	.0895
500-750m.	(.107)	**500-750m.**	(.109)	(.0713)
Ln(tot. SU emp.)	-.291**	**—**	—	—
750-1000m.	(.118)			
Ln(tot. SU emp.)	.0945	**—**	—	—
1000-1250m.	(.144)			
		Stock SU's	.0215**	.0178**
		0-250m.	(.00423)	(.00452)
		Stock SU's	.0142**	.00994**
		250-500m.	(.00250)	(.00283)
		Stock SU's	.00261**	.00380**
		500-750m.	(.00112)	(.00121)
N	164		164	164
Pseudo R^2	—		.50	.51
Sargan [p-value]	2.39 [.793]		—	—
1st stage avg. R^2	??		—	—
avg. F; (min F)	??; (??)			

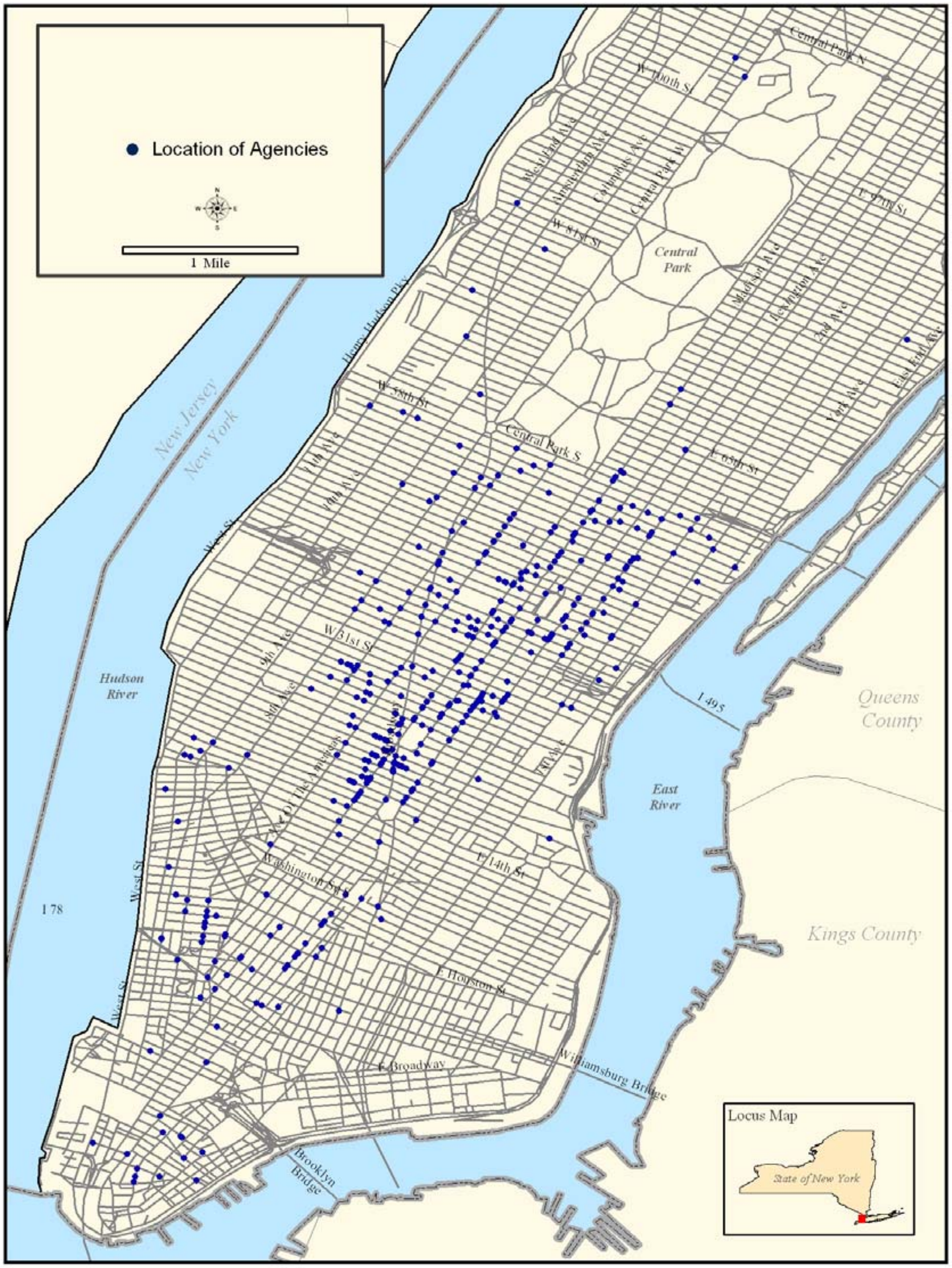
The instruments for column 1 are distance to Rockefeller Center, 1970 log of total housing units, 1970 share of housing units in buildings with less than 5 units, log of total SU advertising agencies employment in each of the five rings in 1977, total count of all establishments in each of the first three rings in 1977, and log of 1992 total number of establishments in the own tract.

Table 8. Heterogeneity: 92-97 Movers

	Logit	Logit	Logit with tract fixed effects	Tract Fixed Effects 2SLS	Tract Fixed Effects OLS
Ln(total no. of	.296	.617		.867**	.872**
establishments)	(.196)	(.496)		(.303)	(.160)
Ln(rent/sq. ft.)	-.0933	-3.91**		-5.35**	-5.31**
	(.633)	(1.82)		(1.04)	(.482)
Stock SU's	.0198**	-.0449*		-.0490**	-.0677**
0-250m.	(.00929)	(.0256)		(.0190)	(.00827)
Stock SU's	.0181**	.0166		-.00297	-.000387
250-500m.	(.00575)	(.0161)		(.0142)	(.00443)
Stock SU's	.00592**	.0193**		.0139**	.0141**
500-750m.	(.00253)	(.00677)		(.00495)	(.00191)
Ln(quality)	.259**	.336		-.320	.125
0-250m.	(.0970)	(.262)		(.271)	(.0841)
Ln(size)*ln(est)		-.0859	-.193		
		(.0895)	(.122)		
Ln(size)*ln(rent)		.748**	1.05**		
		(.313)	(.400)		
Ln(size)*ring 1		.0141**	.0175**		
stock		(.00475)	(.00578)		
Ln(size)*ring 2		-.000430	-.000481		
stock		(.00287)	(.00137)		
Ln(size)*ring 3		-.00247**	-.00220*		
stock		(.00123)	(.00131)		
Ln(size)*ring 1		-.0104	-.0176		
quality		(.0478)	(.0723)		
N [movers]	15580 [95]	13448 [82]	13448 [82]	40	40
R^2		.209	.0774	.828	.826
Sargan [p-value]				1.10 (.89)	

For the second to last column, instruments are distance to Rockefeller Center, 1970 ln(total housing units), share of 1970 units in buildings with less than 5 units, 1977 advertising agency stock for rings 1-3, 1977 total establishments for rings 1-2, ln(median advertising agency payroll size in 1977) for first ring, and 1992 total own tract establishment count.

Figure 1: Locations of Advertising Agencies in Manhattan.



Appendix A. Office Rent Interpolation for NY City

We use the rent for class A office space as an appropriate rent measure for advertising agencies. The data is provided to us by Torto Wheaton Research. It covers 51 ZIP codes in Manhattan in 1992. The data includes the asked price for class A office space in dollar per squared feet. However the data do not provide us with the rent at Census tract level. Thus we interpolate the rent data for tracts in the Southern Manhattan. First, we assign the rent to the centroids of the ZIP codes. Second, we constructed a smoothed rent surface using the spline method (spline estimates values using a mathematical function that minimizes overall surface curvature, resulting in a smooth surface that passes exactly through the input points). Finally we interpolated the rent for a tract by averaging the value of the rent surface over the tract.

Figure A-1. Class A Office Rents for Lower Manhattan ZIP codes.

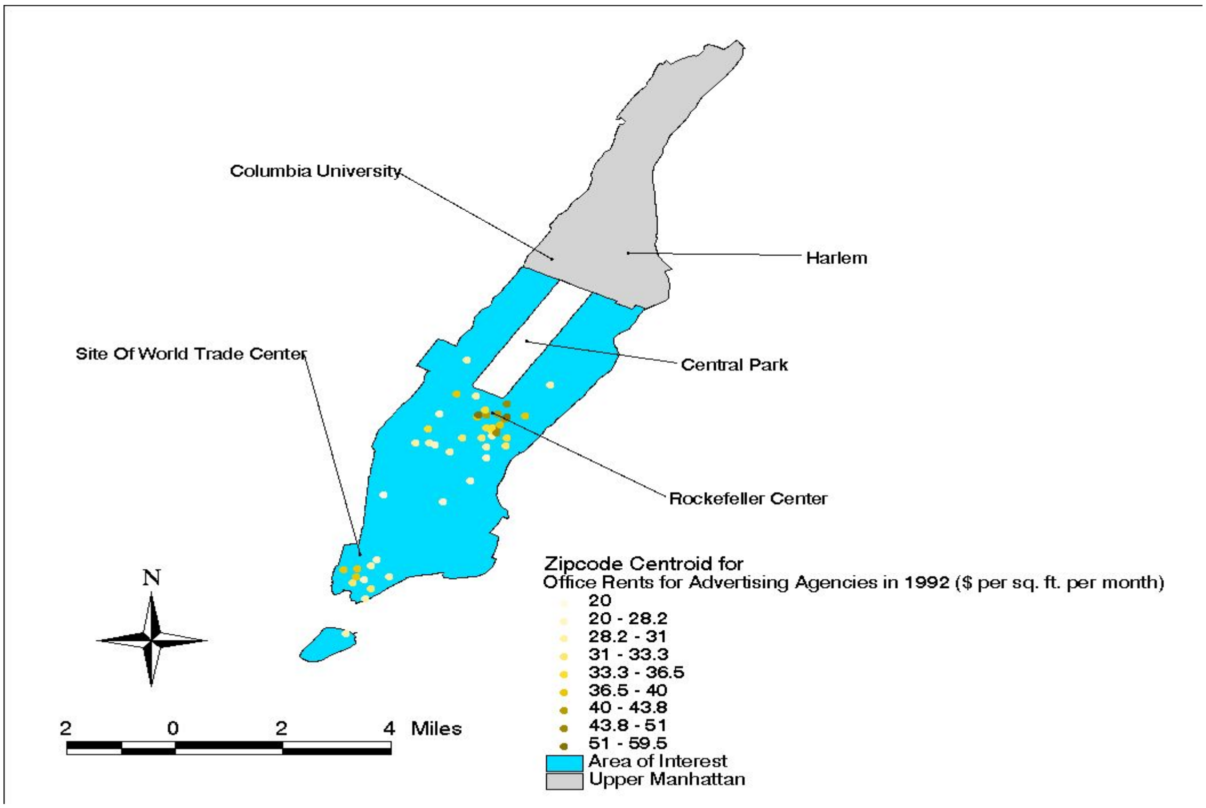
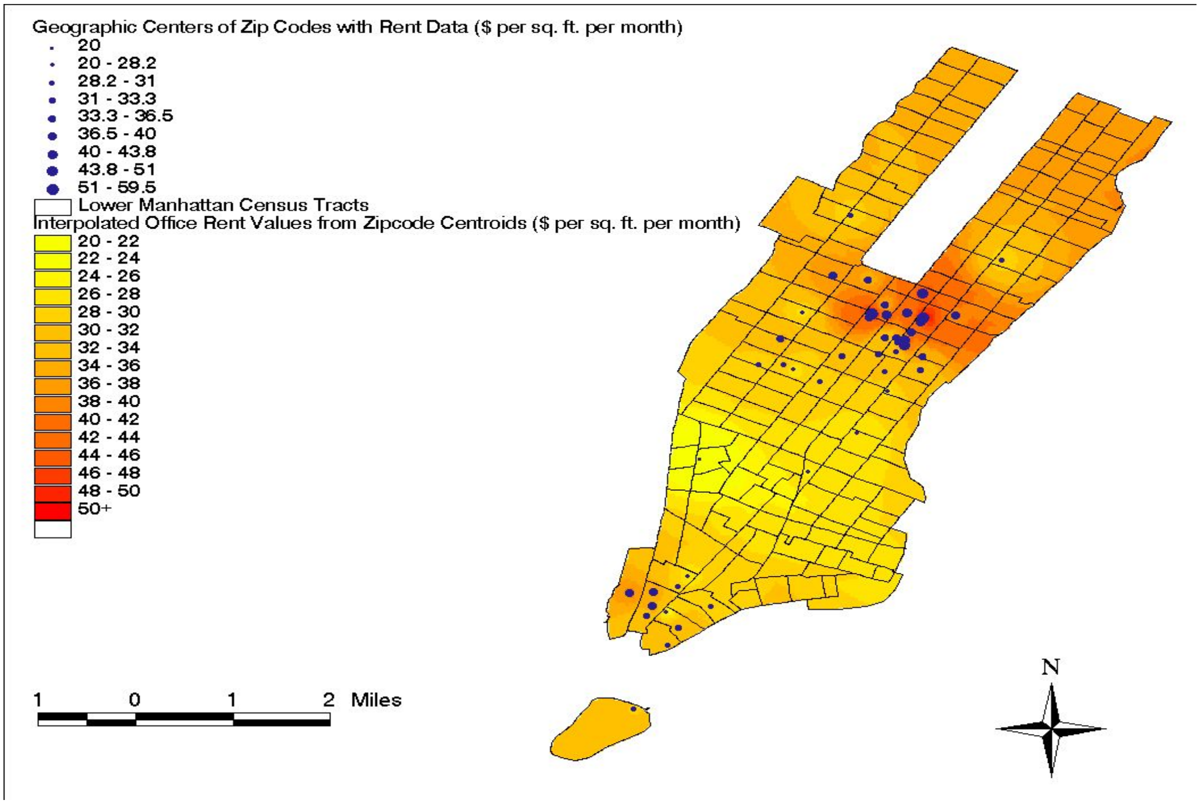


Figure A-2. Class A Office Rent Surface for Lower Manhattan.



Appendix B. Basic Statistics for Covariates and Results for MU's

Table B-1. Means and Standard Deviations of Tract Attributes

Tract Attribute	Raw Mean	Raw S.D.	Weighted Mean	Weighted S.D.
SU Stock 0-250m.	5.80	9.74	17.5	15.8
SU Stock 250-500m.	17.3	20.3	36.8	22.2
SU Stock 500-750m.	37.4	49.0	91.4	69.0
SU Stock 750-1000m.	34.3	34.6	52.6	33.4
SU Stock 1000-1250m	54.3	52.9	96.3	61.4
Ln(total no. establishments)	6.07	1.3	7.1	.88
Ln(rent/sq. ft.)	3.32	.17	3.36	.19
Ln(median payroll of SU's)	3.30	2.73	5.24	1.98
N	164	164	502	502

Attributes are weighted using the number of births in tracts.

Table B-2. Results for MU's

	Ordinary Poisson	Ordinary Poisson	IV estimation
Ln(total no. of	.528**	.576**	.679
establishments)	(.230)	(.0767)	(.574)
Ln(rent/sq. ft.)	-.1.61**	-2.25*	-2.68
	(.807)	(1.28)	(2.98)
Stock MU's <	.262**	.339**	.343**
250 m.	(.0439)	(.0657)	(.111)
Stock MU's	.118**	.166**	.0865
250-500 m.	(.0248)	(.0416)	(.0611)
Stock MU's	.0216	.0307	-.00216
500-750 m.	(.0190)	(.0197)	(.0498)
Stock MU's	-.0461*	-.0465*	.00271
750-1000 m.	(.0257)	(.0260)	(.0696)
Stock MU's	-.0215	-.0198	-.00187
1000-1250 m.	(.0197)	(.0211)	(.0406)
Stock SU's	-.00219	-.0148	
<500 m.	(.00147)	(.00988)	
Broadcast		-.0000547	
Emp.<500 m.		(.0000567)	
Constant	.0228	1.97	2.00
	(3.05)	(4.08)	(12.0)
N [births]	164	164	164
Pseudo Rsq	.460	.468	[.206]
[Sargan p-val]			

Column 1 contains the ordinary Poisson and column 3 the IV Poisson for a baseline specification. There are very strong effects in rings 1 and 2 for ordinary Poissons and in ring 1 for IV results. While for SU's if we increase the number of relevant neighbors in ring 1 by 1 expected births rise by 2%, here for MU's, they rise by 34%. The high coefficients appear to arise from small numbers and an implied non-linearity in scale effects, where the average number of MU's in ring 1 is .8, versus 6 for SU's. In ordinary Poissons, a quadratic for MU's results in significant ring 1 coefficients of .822 and -.0478, so that in moving from 6 to 7 neighbors expected births rise by 2.0%, and positive marginal effects disappear by 9 neighbors. For SU's in a quadratic, if we increase neighbors from their mean number of 6 to 7, expected births rise by 2.8%. Column 2 shows MU's do not benefit from being near SU's. In fact in all formulations we tried such as having 3 or more SU ring variables in the ordinary Poisson or results under IV estimation, we never got positive, even modestly significant results for SU variables.