THE REVIEW OF ECONOMIC STUDIES

Review of Economic Studies (2008) **75**, 1011–1038 0034-6527/08/00411011$02.00

Networking off Madison Avenue

MOHAMMAD ARZAGHI
American University of Sharjah

and

J. VERNON HENDERSON
Brown University

First version received October 2006; *final version accepted December* 2007 (*Eds.*)

This paper studies the advertising agency industry in Manhattan to infer networking benefits among agencies in close spatial proximity. We use economic census data that allow us to distinguish locations at a fine level of geographic detail, so as to infer the strong effect on productivity of having more near advertising agency neighbours. Paying close attention to identification issues, we show, however, that there is extremely rapid spatial decay in the benefits of more near neighbours, even in the close quarters of southern Manhattan, a finding that is new to the literature. This suggests that high density of similar commercial establishments is important in enhancing local productivity for those industries found in large cities, where information sharing plays a critical role. Our results indicate that the benefits of more near neighbours are largely capitalized into rents rather than wages, challenging an existing literature, which estimates wage equations alone to infer agglomeration benefits.

New York County, or Manhattan, accounts for 24% of advertising agency receipts in the U.S. 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 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 for exchange of ideas and expertise among close neighbours is critical to many agencies' success (Arzaghi, 2005). We interpret localized networking as a case of information "spillovers", or Marshall's (1890) "mysteries. . . in the air", but ones involving more deliberate information exchanges.

This paper studies patterns of births of advertising agencies across neighbourhoods in southern Manhattan in a particular time interval to infer the shape of the expected profit function. The paper examines the effect on productivity and profitability of having nearer advertising agency neighbours and hence better opportunities for meetings and information exchange. We have census tract data for Manhattan that allow us to distinguish locations at 250 metre increments. With this fine level of geographic detail, we can infer the high benefits of close interactions. We show that there is extremely rapid spatial decay in the benefits of nearer neighbours 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 industries where information sharing plays a critical role, as modelled in Lucas and Rossi-Hansberg (2002).

Our results indicate that advertising agencies in Manhattan trade-off the higher rent costs of being in bigger clusters closer to the "centres 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, may be largely capitalized into rents rather than wages.

This study differs from and challenges aspects of traditional empirical work on urban scale externalities. First, almost the entire literature, as reviewed in Rosenthal and Strange (2004), examines externalities in manufacturing. However, manufacturing is found disproportionately in smaller and medium size cities, as well as in the rural fringes of cities (Kolko, 2000). If we want to understand the *raison d'être* of larger cities and why certain industries are drawn to high rent, high wage cities like New York with historically built dense neighbourhoods, we need to look at high-end business and financial services such as advertising.

Second, the literature on externalities examines either total factor productivity benefits (*e.g.* Henderson, 2003) or wage gains (*e.g.* Glaeser and Mare, 2001; Rosenthal and Strange, 2006) of greater agglomeration at the aggregate county or metropolitan area level. For wage studies, the problem is that within county, not cross-county, variation in information spillovers may be salient, and spatial variation in the benefits from localized information spillovers may be capitalized into commercial rents, not wages. For advertising agencies, networking benefits are so localized that they are exhausted within the confines of subareas 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 (2006) introduce spatial decay into externality estimation, they look just at manufacturing (and software), look only for wage capitalization effects, and assume no spatial decay within the first mile, whereas, here, effects completely decay within a half mile.

Third, this literature debates what the nature of scale benefits are—information spillovers, labour market externalities, or reduced transactions and transport costs of exchange of intermediate inputs as in the new economic geography (Fujita, Krugman and Venables, 1999) —with no clear resolution. Looking within a county allows us to isolate certain specific sources of externalities. We know what we capture are not labour market externalities: southern Manhattan is all one labour market, drawing workers from all over the New York PMSA and even beyond.[1] We will argue below that, given how advertising services are marketed, what we isolate are proximity benefits for information exchange and networking, rather than some type of shopping centre externality (Dudey, 1990) or market potential effect, although we allow for such effects. We also look for address signalling effects, such as a particular benefit from locating in the Madison Avenue area of Manhattan. In terms of other localized scale externalities outside the own industry, we explore the benefits of access to buyers (*e.g.* headquarters), suppliers (*e.g.* graphic services), and outlets (*e.g.* broadcasters).

Finally, there is the issue of identification of causal effects. Cross-MSA and cross-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 climates, 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); however, a basic identification problem remains. When examining how changes in the local external environment of a firm (*e.g.* numbers of other local firms in the same industry) affect firm productivity in a panel context, such changes are surely not random. They may be driven by changes in unobserved conditions that may also affect both firm productivity and the external environment. The challenge is to find valid instruments for variations in the external environment.

By utilizing a finer spatial scale, the usual county level unobservables are not an issue in this paper: Manhattan is all one legal–regulatory–business climate and infrastructure system. We also have a rich set of observed neighbourhood characteristics as potential controls. However,

1. There is evidence of very modest wage gradients (over longer distances than just within Manhattan) in large metro areas. But it is ascribed to the fact that those living further from the centre earn lower wages because they work locally and have lower commuting costs (supply side) and that clusters further from the city centre may have lower-scale benefits (demand side), rather than there being different labour markets within a metro area (McMillen and Singell, 1992).

there remain neighbourhood unobservables, such as those where 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. These affect current rent prices and location decisions of different economic activities. We argue that these types of unobservable attributes change with time, so unobservable, relevant attributes today are largely uncorrelated with those from, say, 15 years earlier. To try to achieve identification, our strategy is to instrument for current neighbourhood conditions with very particular historical attributes of neighbourhoods, which remain correlated with covariates but not with current error drawings. We will present evidence that neighbourhood characteristics change over time and that particular historical attributes are correlated with current covariates but not error terms. Since such a strategy raises issues, we devote considerable space to evaluating identification.

In the first section, we examine key aspects of the advertising agency industry, both to give a sense of why advertising is such a relevant industry in examining urban scale externalities and to identify key aspects of the industry in Manhattan relevant to estimation. In Section 2, we specify a model of location choice within Manhattan for single-unit (SU) advertising agencies, to help us think about empirical issues. Then we turn to econometric implementation and identification and results. Section 3 gives our basic results and specification tests. Section 4 considers a variety of extensions, mainly examining robustness of the basic results and considering heterogeneity of effects. Section 5 concludes.

1. ADVERTISING AGENCIES

This section gives relevant background information on the advertising agency industry, in general, and in Manhattan.

1.1. *The nature of advertising agencies and networking*

In 1997, advertising expenditures in the USA totalled \$187bn, about 2·2% of GDP. Of this, only \$17bn involves advertising agency receipts (NIPA, BEA).[2] However, of the \$17bn, \$10bn comes from commissions on \$89bn 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. Much of the rest of advertising agency income is associated with what 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 high-skilled labour intensive services for a client—market research, creation of an advertising campaign, production, media placement, and client management. Advertising agencies derive a significant portion of their income from fees and billings for creative services rendered in the design of campaigns. In addition, a successful creative design can win an account and generate media commissions. Like R&D (Jaffe, Trajtenberg and Henderson, 1993), where creativity

2. Advertising is split about 55–45 between national vs. 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 for 32%, and yellow pages and direct mail for 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.

plays a central role, information sharing and information diffusion are thought to be critical to an agency's success.

Networking appears to operate on several levels. One occurs at the design stage of advertising campaigns where an agency receives a "request for proposal" (RFP) for an advertising campaign and seeks help from its network in designing the best proposal (Arzaghi, 2005). Another occurs in the exchange of information about suppliers, outlets, and market conditions in the home market of clients (Hameroff, 1998).[3] Finally, firms utilize materials and research from other firms in its network, and as Economic Census data indicate below, they collect fees from clients for these materials, which they remit to the firms from which they received the materials (after deducting a commission). Clustering facilitates localized networking, where repeated face-to-face contacts are critical.

These assessments are partly based on our interviews of executives of a dozen advertising agencies in Manhattan and Chicago. We asked about their contacts with advertising professionals outside of the agency. First, the interviews indicate that telephone and email are the most frequent means of contact. However, all interviewees emphasized that every such contact is contingent upon initial (and ongoing) face-to-face meetings. In fact, for these agencies, the majority reported that their last contact was in the form of a face-to-face meeting.

Second, the executives said that their main goals in contacts are to supplement their limited in-house capacity, in terms of gathering both ideas in preparing proposals and sufficient materials and labour to fulfil a particular contract. As a simple example of the latter, agency A received work to redesign a set of presentation slides for a client. The people in agency A worked on the set of slides for a week and presented a sample to the client. The client was happy with the sample. Then the agency learnt that the work involved not only the 100 pages in the set of slides discussed in the initial meeting, but also that there were 10 other similar cases that needed to be done in about 10 days. This was beyond the capacity of the agency. To help keep the account, the head of the agency A utilized a contact in agency B he trusted could help with the job. That contact was currently two blocks away. They have been involved in a business relationship that started 10 years earlier. Although both had moved in the last 10 years, their very first contact involved collaboration on a small project, when they were both located also in close proximity to one another in Manhattan. Other examples of lasting cooperation with firms in close spatial proximity were given in the interviews.

As an example of how contacts matter in a response to an RFP, advertising agency C in Manhattan received an RFP from an advertiser, which recently had closed its account with another agency nearby in Manhattan, D. The executive of C had a close contact in agency D, which lost the account. Even though the executive in C was limited in what he/she could either ask the contact in D or report to us as interviewers about aspects of the client's business (confidentiality constraints), he/she reported consulting with the friend in D about personalities and tastes of relevant people in the advertiser firm.

1.2. *Aspects of the structure of the industry*

The notion that networking is relatively important to advertising agencies 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 to business services more generally, noting how business services are much less concentrated than manufacturing. Establishments

3. 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 11 huge national and international formal networks of larger agencies. These formal networks primarily exchange information about marketing conditions across cities and countries.

TABLE 1

1992 Economic Census numbers on concentration in advertising†

	Advertising agencies		All other business services	
	Share of total (%)	No. units	Share of total (%)	No. units
Counts of establishments with under five employees	58	6890	52	133,380
Sales by establishments with under five 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 firms	90	12,453	81	236,212
Sales by single-unit firms	55	12,453	45	236,212

†The establishment numbers cover just operating establishments. 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 U.S. establishments	7·3%
Share of U.S. employment	20%
Share of U.S. receipts	24%
Share of media billing	31%

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 21 in the rest of business services. SU firms account for 90% of establishments and 55% of sales in the advertising agency industry, compared to 81% and 45% in other business services.

While the advertising industry is non-concentrated in terms of firm size, it is heavily spatially concentrated in New York County, or Manhattan, as shown in Table 2. In 1997, Manhattan had 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 U.S.; 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. In line with this assessment, we note that Manhattan's share of headquarters has declined significantly over time and now is only 3% of total headquarters employment in the U.S.

In the empirical work, in examining the advertising agency business, we focus on the activities of single-unit (SU) firms, as opposed to establishments of multi-unit (MU) firms, with the latter case analysed in Appendix B. MUs are large firms with multiple, very large establishments, which seem to serve different types of clients. In New York, SUs have an average size of 11 employees, while MU *establishments* (not firms) have an average size of 155 employees. MUs are much more reliant on income from media receipts, indicating their greater role as intermediaries between advertisers and the media. In 1997 in New York, they got 48% of receipts from media billings while SUs got 28%. In contrast, SUs are more involved in the creative design of products, with much more income from fees for services, and as noted above, they seem to engage much more in networking. In 1997, they received 21% of their income from commissions and fees for services provided by *other* agencies, compared to just 6% for MUs, an indication of greater networking, or "sharing" through cross-firm indirect sales by SUs. Examples of such cross-agency services cited by the census questionnaire include artwork, plates, and marketing and other research.

Why don't we combine MU and SUs, or focus more on MUs rather than relegating those results to an appendix? First, in the text and Appendix B, we argue that firms in each group seem to interact within their own group (SU and MU), but not across groups. In Pearson χ^2 tests of similarity of location patterns discussed later, it is clear that MU and SU establishment stocks have different location patterns across zip codes in Manhattan in all years of our data. This lack of interaction presumably relates to our inference that they operate in different markets. Second, sample sizes for MUs are 10% of those for SUs. In pooling, SU effects dominate all results (*i.e.* we get basically the same results as for SUs) and making detailed inferences separately for MUs is more tenuous because of limited sample size. MUs are also less interesting for our purposes. Cross-firm networking is less important for MUs for their more media-based activities, and our interviews suggest MUs may rely more on intra-firm networks across their own establishments.

Before proceeding, we note there are a whole set of industrial organization questions we do not address because of data limitations, as well as what is feasible in the scope of one paper. We look at location patterns of births of new firms to infer what neighbourhood attributes they value. There is a dynamic of new firms being split-offs where an employee leaves an existing firm to start a new firm and then a later dynamic of deaths of firms, which include buy-outs. We do not have information on this process, nor do we examine (or observe) how multi-unit firms arise through acquisitions vs. creation of new establishments, nor the segmentation of the market between single and multi-unit firms. We focus only on networking benefits among firms to document the nature and speed of spatial decay of such benefits.

1.3. *Manhattan*

Within Manhattan, advertising agencies are closely clustered. Because of disclosure criteria, we are not allowed to use census data to map agency locations. Although the empirical work to follow is based on the entire population of advertising agencies in the Census Business Register (CBR or Standard Statistical Establishment List, SSEL), we use The Advertising Red Books online for 2003 to generate maps. The Red Books give the location of about two-thirds of the agencies in Manhattan, usually larger agencies.[4] Figure 1 shows these agency locations in the southern half of Manhattan, defined as the area below 90-th Street. Almost all agencies in New York are located below 90-th Street. 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 break out single and multi-unit firms, so we cannot map them separately. If we compare Figure 1 with the rent gradient map in Figure A2, 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, although the Red Books probably underrepresent smaller agencies in fringe locations. We note for 1992 in the census data, the simple correlation between the number of single-unit advertising agencies in a tract and the log of rent price is 0·23.

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.* stocks in 1977 vs. 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 from the census in all years (1977, 1982, 1987, 1992, 1997).

4. This is a publication of National Register Publishing. The Red Book includes self-reported data.

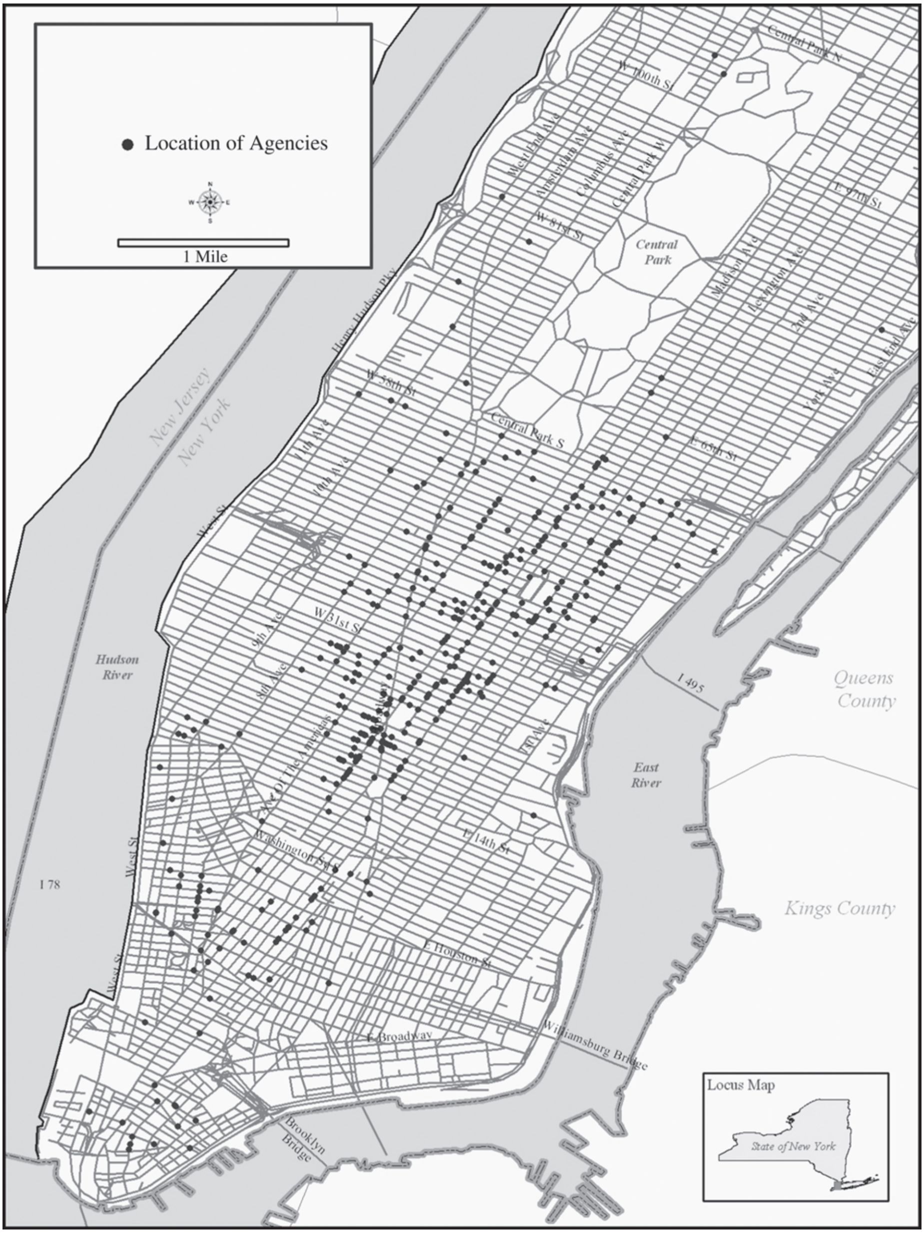

FIGURE 1

Locations of advertising agencies in Manhattan

Using Pearson χ^2 tests[5] to compare location patterns across zip codes for all of Manhattan, we examine a key aspect of location patterns, relevant in the sections to follow. Location patterns of single-unit advertising agencies in Manhattan have changed over time. Specifically:

(1) Comparing 1987 (or 1982) with either 1992 or 1997, the stock location pattern of single-unit firms in 1987 (or 1982) differs from the pattern in 1992 or 1997.
(2) Corresponding to (1), births in 1982–1987 vs. 1992–1997 have different spatial distributions.[6]

Patterns change from 1987 to 1992 reputedly in response to relative rental cost changes across areas of Manhattan at that time, moving advertising agencies out of some of the highest rent areas. Such relative rent cost changes must be correlated with and driven by changing tract characteristics over time. Both the changing patterns of advertising agencies' locations and the relation to changing rents will be noted below when discussing identification. Most births of single-unit firms since 1992 have happened in medium and lower-rent areas and hence off Madison Avenue.

2. NETWORKING, LOCATION CHOICE, AND THE EMPIRICAL SPECIFICATION

In this section we briefly present a model of the networking benefits for a firm, which leads directly to an empirical specification. We discuss data and estimation issues, focusing on an identification strategy.

2.1. *Formulating local network contacts*

In the empirical work we will infer the productivity benefits to new firms of locating in neighbourhoods with more advertising agencies with which to network. Although our estimation is not strictly structural, we provide a simple framework to motivate our thinking, specification, and interpretation of results for those who like a more structural interpretation. While the specification and jargon is informed by the large theoretical literature on networks (*e.g.* Jackson and Wolinsky, 1996; Bala and Goyal, 2000; Goyal and Moraga-Gonzalez, 2001), 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 pairwise extent of contact, independent of other pairwise arrangements.

Networking can involve social meetings with employees of other agencies, drop-in visits at other firms, scheduled business meetings where team members of different firms get together, or contacts through intermediaries, such as local graphics firms or contract designers serving several agencies. The distance between agencies plays a key role in the extent of networking, representing commuting time for face-to-face meetings or costs of monitoring of one another's activities. In the model, 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.

5. The Pearson test statistic for homogeneity is $T = \sum_{i=1}^{r}\sum_{j=1}^{c}\frac{(m_{ij}-e_{ij})^2}{e_{ij}}$ where $e_{ij} = Np_i p_j \equiv N\frac{n_{i.}}{N}\frac{n_{.j}}{N}$. Respectively, $p_i = n_{i.}/N$ and $p_j = n_{.j}/N$ are the ML estimates of the marginal distributions for rows and columns, where $n_{i.}$ is the sum of the counts across the columns (*e.g.* the sum of the agency stocks in location i in 1987 and 1997) and $n_{.j}$ is the sum across rows (*e.g.* total number of agencies in 1987). m_{ij} is the count in each cell. T is distributed χ^2 with degree of freedom $(r-1)\bullet(c-1)$. The degrees of freedom are $(r-1)\bullet(c-1) = rc-c-r+1$, where the intuition is that, in estimating marginal probabilities from the initial rc degree of freedom we "use" 1 to calculate the total sum of the counts and $r-1$ for the row and $c-1$ for the column marginal counts. Details of the test and degrees of freedom are in Read and Cressie (1988), p. 28, as well as Conover (1999), p. 240.

6. These differences apply to establishments of MU firms as well; but as we already noted above, their location patterns differ from those of SU firms.

Consider a firm operating at location j. First we look just at the portion of the specification of its technology that relates to networking with one other firm. The value of a contact of firm j with a firm at location i is represented by

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

In (1), V_{ij} is the volume of communications between the firm at location j and a firm at i, where we assume that $0 < \gamma < 1$, so there are decreasing returns to volume in any pair of contacts. A "match" parameter between firms, u_{ij}, is drawn on the interval [0,1], which is specific to the match between the firm at j and that at i, resulting in variation in the extent of contacts across firms. The cost per unit of communication, $c(\delta_{ij})$, is an increasing function of the distance from i to j. Total communication costs comprise firm labour costs; one can define V_{ij} as labour units involved in communication, with $c(\delta_{ij})$ representing lost work time from employees having to travel further to interact with another firm. F is the initial fixed cost of making a contact with a firm at another location, upon which u_{ij} is learnt.[7]

Once the firm spends F to learn u_{ij}, it chooses V_{ij} so that $V_{ij}^* = (\gamma/c(\delta_{ij}))^{1/1-\gamma}(u_{ij})^{1/1-\gamma}$. If $c(\delta_{ij}) = c(\delta_{ji})$ and $u_{ij} = u_{ji}$, then $V_{ij}^* = V_{ji}^*$ and communications are reciprocal and two sided. Given V_{ij}^*, equation (1) may be written in indirect form, where the *ex post* value of contact with any firm is $(1-\gamma)\gamma^{\gamma/(1-\gamma)}c(\delta_{ij})^{-\gamma(1-\gamma)}(u_{ij})^{1/(1-\gamma)} - F$. Note the volume of communication and its value are decreasing in distance and communication costs. The value of *ex post* profits suggests there is a maximum distance, $\delta_{\max}$, beyond which $c(\delta_{ij})$ is high enough relative to F so that, *ex ante*, a firm seeks no contacts with firms beyond $\delta_{\max}$, as we will see next.

In empirical implementation, we take this specification of the value to firm j of communications with one other firm and imbed it in a technology, which allows for more firms and a choice of office space given rent, all designed to yield a simple estimating equation. In the specification, l_j is the amount of space consumed by firm j at a unit rental cost, R, in the location under consideration, and A_j represents amenities for firm j at the location. These amenities include observed components from both the own and potentially surrounding locations (*e.g.* neighbourhood commercial scale and access to broadcasters),[8] as well as unobserved components in the error structure given below.

In the overall specification of technology of a firm, *ex post* profits are

$$A_j l_j^{\alpha} \exp\left(\sum_{i}^{\delta_{\max}} \{[V_{ij}^{\gamma} u_{ij} - V_{ij}\mathrm{c}(\delta_{ij})] - F\}\right) - R_j l_j. \tag{2}$$

Optimizing with respect to choice of V_{ij} gives *ex post* profits as

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

First, we calculate expected values (assuming draws of firm-to-firm matching parameters, u_{ij}, are i.i.d.) for the firm at location j, maximize *ex ante* profits with respect to l_j, and substitute back in for the optimal value of l_j. Second, given the discrete nature of our geographic units (tracts), we will distinguish discrete locations relevant to the firm at j as discussed below and

7. While u_{ij} can take the value 0 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.

8. We examine just the net effects of these. We could, of course, specify a similar structure for other activities—benefits and costs of access, which would just follow what we are doing for access to other advertising agencies.

allow there to be n_i firms at each such location. Taking logs, after rearrangement, expected profits from choosing location j are

$$\tilde{\pi}_{\mathrm{j}} \equiv \ln \Pi_j = C_j - \frac{\alpha}{1-\alpha}\ln R_j + \frac{1}{1-\alpha}\sum_i^{\delta_{\max}} n_i \ln \mathrm{E}\left[\exp\{(1-\gamma)\gamma^{\frac{\gamma}{1-\gamma}} c(\delta_{ij})^{\frac{-\gamma}{1-\gamma}} u^{\frac{1}{1-\gamma}} - F\}\right]. \quad (3)$$

For the last term in equation (3), the empirics estimate the effect of an increase in the number of firms at location i on the profits of a firm located at j. Thus we estimate a set of coefficients that we expect to decline with distance δ_{ij}, and match the expression $\ln \mathrm{E}[\exp\{(1-\gamma)\gamma^{\gamma/1-\gamma} c(\delta_{ij})^{-\gamma/1-\gamma} u^{1/1-\gamma} - F]/(1-\alpha)$ in (3). In estimation, we look for the $\delta_{\max}$ for the firm at j, beyond which it does not pay to try to communicate with firms in those locations because $\ln E[\exp\{(1-\gamma)\gamma^{\gamma/1-\gamma} c(\delta_{ij})^{-\gamma/1-\gamma} u^{1/1-\gamma} - F] < 0$. Our estimation is not strictly structural; for that, we would need to assume a specific PDF for u_{ij} and shape of $c(\delta_{ij})$, for which we do not have good priors. Moreover, the model is meant to motivate the empirical formulation and interpretation, rather than impose one structure over others. Finally in (3), C_j contains a multiple of the original A_j, as well as a collection of parameters, C. In particular, we assume $C_j = C + \tilde{C}_j + \eta_j$ where $\tilde{C}_j$ represents the effect of neighbourhood characteristics and η_j is a firm idiosyncratic "taste" (in terms of the idiosyncratic effect on profits) for tract j drawn from an extreme value type I distribution. Therefore, we can rewrite expected profits as $\tilde{\pi}_j = \pi_j + \eta_j$, where π_j is log expected profits of a representative firm if it locates in tract j.

2.2. *Empirical implementation and data*

The model we estimate is based on equation (3). In the treatment of space, locations are census tracts, which in New York are at a finer level of spatial detail than zip codes. Moreover, a census tract is intended to represent a true micro "neighbourhood", as most tracts in Manhattan coincide with the city blocks between two avenues and a few streets. The main variables constructed for each census tract are: an interpolated rental cost per square foot of class A office space, availability of commercial sites measured by the stock of private establishments, distance measures to other advertising agencies, and distance measures to other activities such as broadcasting establishments (where ads are placed), headquarters (clients), or graphic services (inputs), as discussed below.

2.2.1. Empirical formulation. In estimation we look at the location choices of new firms in Manhattan, where in the specification, the given number of births, which occur in Manhattan in a particular time interval allocate themselves across census tracts to maximize profits. In viewing the choice of tracts, a firm looks at each tract as one choice with, as noted, a firm-tract specific draw, η_j, from an extreme value type I distribution. These firm-tract-specific shocks to profits are the reason in the usual logit framework why different firms are spread across different tracts (as opposed to all choosing the same tract). The probability that a firm chooses tract j is

$$P_j = \frac{\exp(\pi_j)}{\sum_i \exp(\pi_i)}. \quad (4)$$

One of the covariates (in $\tilde{C}_j$) used in all formulations is the log of the total number of private establishments in the neighbourhood, m_j. This covariate could be viewed as capturing localized general agglomeration economies from having more commercial activities in a firm's own neighbourhood, including any general market potential or shopping centre externalities, beyond the specifics of headquarter locations in Manhattan, which are treated separately. However, as noted in Section 1, 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 do not expect market potential effects or shopping centre externalities within Manhattan. Nonetheless, controls on commercial scale of the own tract, as well as potentially neighbouring tracts, allow for such effects.

Holmes (2005) suggests a second role for this variable, which measures total commercial spaces of a neighbourhood, that is based on a reformulation of the error structure. 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 on location possibilities within a tract. In that case, the coefficient of m_j could be interpreted as $1-\sigma$, where σ is approximately the within tract correlation of error drawings in a nested logit framework. Under this interpretation, equation (4) 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).

For births, all covariates are tract characteristics, since we observe no firm characteristics. In this case, if the expected number of births in a neighbourhood, λ_j, takes the usual form $\exp(\pi_j)$, the problem may be equivalently estimated as a standard Poisson model of birth counts per tract (Guimarães, Figueirdo and Woodward, 2003). We start by presenting ordinary Poisson results. However, a potential problem in estimation is that there may be unobserved tract characteristics in $\tilde{C}_j$. These omitted characteristics, such as trendy eating and socializing spots and construction sites affect profits and also values of other covariates, such as the stock of other advertising agencies with which the new firm might interact. Thus, error draws may include unobserved neighbourhood characteristics that are not firm specific, which affect profits of all firms, and are correlated with observed neighbourhood covariates. As such, we want to alter our estimation strategy and also identify instruments that are correlated with neighbourhood covariates, but are exogenous to contemporaneous error draws. Assuming such instruments exist, we formulate a moment condition, where using semi-parametric methods relaxes the restrictive assumption of the Poisson model that mean and variance are equal and avoids over- or underdispersion problems.

If we define

$$\upsilon_j \equiv B_j - \lambda_j = B_j - \exp(\pi_j),$$

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

$$E[\upsilon_j \mid Z_j] = 0, \tag{5}$$

where Z_j are instruments. We use non-linear generalized method of moments (GMM) applied to this one moment as suggested in Windmeijer and Silva (1997) and Mullahy (1997) to estimate the parameters. Of course, if Z_j are simply the covariates of the model, estimation of (5) gives identical coefficients to those in the Poisson and logit.

2.2.2. Data and covariates. Data are from the CBR (SSEL), which covers all establishments in the U.S. We analyse the location decisions of births in Manhattan from 1992 to 1997.

TABLE 3
New York City births

	1997 Stock	1997 Stock, born before		1992–1997 Births
		1973	1988	
Single-unit estimate	1088	104	357	545

A birth is a new firm to Manhattan—a firm that did not exist in the 1992 SSEL, but did in 1997. We use birth patterns rather than stock patterns, since births are sensitive to current neighbourhood conditions rather than past ones. We chose the SSEL in Economic Census years because the records on existence and location are updated and most accurate for those years and the Economic Census reports an actual year of birth. We examine births in the 164 census tracts south of 90-th Street, which cover 97·4% of all SU advertising agencies in Manhattan. We exclude the 132 tracts north of 90-th Street because generally they are not viable locations for advertising agencies as well as other class A office space users. As such, also we cannot accurately infer rent data per tract for them.

Total births in Manhattan are given in Table 3. For our tracts south of 90-th Street, the totals are slightly lower, with the 1992 stock being 949 and 1992–1997 births being 502. 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 births for the 148 establishments that in 1997 report their birth year as 1993 or 1994.[9] We first show that results for the two samples are the same and then focus on 1992–1997 births because tract births are more numerous and the count numbers are less noisy.

In terms of covariates, 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 office rents vary threefold within southern Manhattan.

For advertising agency neighbours, for each tract, we construct measures of access of that tract to nearby advertising agencies, by allocating neighbour agencies in 1992 into rings. We define five rings moving out in increments of 250 metres, based on the centroids of tracts. Ring 1 is the count of existing SU advertising agencies in census tracts whose centroid is within 250 metres of the own census tract centroid. Generally ring 1 is just the own census tract. Ring 2 is the count of agencies in census tracts whose centroids are between 250 and 500 metres of the own tract centroid, and so on for rings 3–5, up to 1250 metres. Each additional ring typically contains 3–4 more tracts.[10] We also experiment with other ring divisions, as reported in Arzaghi and Henderson (2006) and discussed later. Finally, we calculate the total number of commercial establishments for each tract and related rings in Manhattan, as well as a variety of other specific commercial activities of relevance to the advertising business as noted earlier.

2.2.3. Identification and instruments. By focusing just on Manhattan, unobserved neighbourhood characteristics include trendy places, construction, security, and neighbourhood

9. Yearly SSEL records for 1993 and 1994 show almost no births, reflecting the fact that birth records are not updated until census years. This is why we limit our data to the Economic Census years.

10. Note that southern Manhattan is about 2600 metres across and about 11,000 metres long.

public services for firms, which change over time. As such, we choose particular historical variables from the 1970 Population Census and 1977 SSEL as instruments for 1992 covariates. At a minimum, this should mitigate the omitted variables bias. Later in the paper, we report many tests, which suggest our assumption that instruments are not correlated with current error terms is valid. Here, we first address what variables are endogenous and what are possible strong instruments.

In the empirical work, the coefficient subject to the greatest bias will turn out to be rents. The effects of current neighbourhood unobservables on businesses in general are capitalized into rents. Better conditions are reflected in higher rents, biasing the rent coefficient upwards and underestimating the negative effect of rents. 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 five or fewer units. These variables reflect the potential over the long term, despite regulations, to convert residences into 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 centre 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 are also potentially endogenous variables. Trendy local conditions not only influence rents and births today, but 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 later in the paper we report test results, which suggest spatial correlation is not a compelling issue, we still instrument for all ring stock variables. For instruments, we use 1977 values of ring stocks.[11]

Use of these instruments raises interrelated issues. If tract unobservables and location patterns of advertising agency stocks do not change over time, use of historical stocks to instrument for current stocks is not valid. Location patterns replicate over time with births replacing deaths by tract and with instruments being correlated with current unobservables. We have already documented that location patterns did change from the 1980's to the 1990's. The driving force cited is relative changes in rent patterns, but such rent changes also imply underlying changes in tract characteristics. But if that is correct, then why are historical stocks strong instruments for current stocks? While there are distinct shifts in location patterns, within historical clusters there is also some degree of persistence in the location of older firms, driven by moving costs. These include physical moving costs, loss of address recognition, and the costs of leaving (defecting from) an existing network (*i.e.* coordination failure of moving clusters en masse). In general, 1977 ring stocks are reasonably strong instruments.

Despite this argument, one may remain suspicious, *a priori*. Besides formal and informal specification tests discussed in Section 3, we conducted many experiments. These include (a) adding various measures of other tract characteristics that could affect location choices, (b) looking at patterns of deaths to show birth patterns are not replicating death patterns, and (c) trying an alternative instrument for the own ring stock.

For the last, an alternative instrument is based on a dramatic change in the way headquarters, as consumers of advertising agency services, 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 0·82. By 1992, the corresponding correlation coefficient had fallen to 0·28.[12] With

11. Actually, we can only assign about 70% of establishments in 1977 to a 1992 census tract.

12. Note patterns of residential vs. commercial locations across Manhattan generate positive correlation between almost any two business activities.

the Internet, cheap air travel, cheap telephone service, and the like, the advertising agency–headquarter interaction has changed. Now, most of their interactions are not face to face. When such interactions are important as in the highly intense design stage of advertising campaigns, agencies usually fly a team to the headquarter location away from Manhattan for some weeks. Given the historical clustering of advertising agencies around headquarters in Manhattan and some persistence in locations of older advertising agencies, the correlation coefficient between 1992 SU advertising agency location patterns and 1977 headquarter location patterns is 0·56, double the coefficient with contemporaneous headquarter location patterns. This makes 1977 headquarter counts by tract a potentially good alternative instrument for 1992 stocks. That is especially the case if one believes headquarters' location choices in 1977 were little influenced by where advertising agencies located, being more dependent on tract characteristics not relevant to advertising agencies like access to financial services.

3. BASIC RESULTS

In this section we start with results for the basic birth specification in Table 4. In that specification, coefficients on included covariates are not sensitive to the addition of other covariates that we experiment with in the next section. The discussion of Table 4 is divided into three parts. First we discuss the effects of IV vs. ordinary Poisson estimations; then we turn to the numerical interpretation of coefficients, and finally we discuss in great detail the validity of our instrumental variable strategy. In Table 5 we examine further identification issues. In Section 4, Tables 6 and 7 explore robustness of results to changes in specification and inclusion of other covariates.

Before presenting details, we summarize the basic results. Rents have strong negative effects on agencies' location decisions and inferred profits. Close proximity of other advertising agencies strongly affects location choices, but such effects dissipate rapidly with distance. That is, firms are willing to pay substantial rents to locate near the centres of action for advertising agencies, as opposed to the fringes. In general, the only other covariate that is important to control for in estimation is total commercial activity in the own tract. Results are very robust to alternative specifications and a variety of controls.

3.1. *IV vs. non-IV results*

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; the next two columns are for 1992–1997 births. For 1993 and 1994 births, column 1 contains ordinary Poisson results (identical to ordinary conditional logit results); column 2 contains the results of non-linear GMM IV estimation (referred to as IV-GMM in the tables). Comparing the two columns, the impact of IV estimation is concentrated on the rent and own (0–250 metres) ring stock coefficients, with anticipated effects. First, the absolute value of the IV coefficient on rent is much higher than the ordinary Poisson coefficient. Higher rents, aside from meaning higher prices (negative effect), reflect better unobserved contemporaneous neighbourhood amenities (positive effect) in the ordinary Poisson specification, biasing the coefficient towards 0. Second, the IV estimation of the own ring stock coefficient (0·0218) is lower (by 35%) than the ordinary Poisson estimation. A larger number of agencies in the first ring may reflect higher unobserved and somewhat persistent—from before 1992—neighbourhood amenities affecting both new births and stocks in 1992, so ordinary Poisson estimates are biased upwards.

We also note that, while the own ring stock coefficient is lower under IV estimation, coefficients for rings 2 and 3 tend to be higher. Why might this be? Later we report results from Moran I tests where we cannot reject the hypothesis that error terms are spatially uncorrelated, but the

TABLE 4

Birth models

	Birth 1993–1994 Poisson (1)	Birth 1993–1994 IV-GMM (2)	Birth 1992–1997 Poisson (3)	Birth 1992–1997 IV-GMM (4)
Ln(total no. establishments)	0·411**	0·530**	0·462**	0·498**
	(0·137)	(0·216)	(0·0756)	(0·117)
Ln(rent/sq. ft.)	−1·72**	−2·74*	−0·819**	−2·57**
	(0·572)	(1·46)	(0·289)	(0·773)
Stock of agencies 0–250 m.	0·0334**	0·0209*	0·0206**	0·0198**
	(0·00817)	(0·0119)	(0·00430)	(0·00758)
[Willingness-to-pay/sq. ft. (%)]	[1·9]	[0·76]	[2·5]	[0·77]
Stock of agencies 250–500 m.	0·0147**	0·0191**	0·0167**	0·0228**
	(0·00494)	(0·00953)	(0·00264)	(0·00538)
Stock of agencies 500–750 m.	0·00451*	0·00648	0·00387**	0·00419
	(0·00250)	(0·00500)	(0·00137)	(0·00259)
Stock of agencies 750–1000 m.	0·000702	−0·00336	−0·00149	−0·00149
	(0·00281)	(0·00639)	(0·00158)	(0·00274)
Stock of agencies 1000–1250 m.	−0·00577**	−0·00519	−0·00219	−0·00241
	(0·00272)	(0·00605)	(0·00147)	(0·00251)
N	164	164	164	164
Pseudo R^2	0·343	—	0·504	—
Sargan (p-value)	—	3·88 (0·694)	—	1·71 (0·945)
First-stage average R^2	—	0·50	—	0·50
Average F (min F)	—	13·0 (7·9)	—	13·0 (7·9)

Notes: The instrument list is distance to Rockefeller Center, 1970 log of total housing units, 1970 share of housing units in buildings with less than five units, stocks of single-unit 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. **Significant at 5% level; *significant at 10% level.

non-rejection is weak. If error terms display some degree of spatial correlation, that could help explain the opposing effects in own ring vs. ring 2 with IV estimation. If own ring and neighbouring ring unobservables improve together, the improvement in the neighbouring ring may draw births away from the own ring, understating the pure networking effect on own tract births offered by ring 2.

Columns 3 and 4 show the effects of IV estimation for 1992–1997 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 (Train, 2000), coefficients are not absolute profit function coefficients as discussed next; the degree of bias in terms of willingness-to-pay rent to have more agencies in ring 1 is larger for 1992–1997 births than 1993–1994 births: a 3·2 (2·5/0·77)-fold increase compared to a 2·4 (1·9/0·76)-fold increase. Third, the IV estimates of the own ring, rent coefficients, and willingness to pay are basically the same for 1992–1997 births and 1993–1994 births. In general, we rely on 1992–1997 results because the sample of births is larger, but the results for the two samples are always very similar.

3.2. *Interpreting 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 indicate the percentage by which a change in a covariate raises or lowers the expected number of births in a tract.[13] To evaluate relevant magnitudes, Table B1 reports means and S.D.s of covariates. Second, we can monetize effects on profits utilizing the rent coefficient. For each ring stock magnitude, we calculate what percentage increase in rent per square foot of office space a firm would be willing to pay to have one more neighbour; or we can ask how many S.D. increases in unit rent a firm would pay for a one S.D. increase in neighbours.

In column 4 of Table 4, a 1% increase in rents leads 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 S.D. change (0·17) in the (log) 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 neighbours, column 4 shows that, as we move from the first ring out in 250-metre increments, the coefficients are 0·020, 0·023, 0·0042, and then 0 beyond 750 metres. Ring 1 and 2 effects are always similar, but at 500 metres, 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 at ring 3. For the hypothesized networking effects, close spatial proximity is critical. Based on numbers of advertising agencies per ring in Appendix B and changes in ring areas as we move out, the number of advertising agencies per unit spatial area does not decline as we move to the 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 metres, 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 is why the ring 1 effect does not dominate the ring 2 effect? First, it could be that commuting costs are non-linear in distance: people may be indifferent between a five minute and 12 minute walk, but not between those and a 20 minute one. Second, although we do not think there are shopping centre or market potential externalities, there may still be a poaching problem. For a firm, while it may be great to have a network member a 5–10 minute walk down the block, having an agency right next door could advertise a neighbour's features to current clients who may choose to visit the agency occasionally.

In terms of magnitudes, if neighbours in ring 1 increase by 1, the expected number of births increases by 2%. For typical variations, the S.D. of ring 1 neighbours is 10.[14] For example, if we increase the number of immediate neighbours by one S.D., that raises the expected number of births by 20%, a large effect. In terms of willingness to pay greater rental prices for more neighbours, column 4 indicates (the number given in square brackets) that a firm would be willing to pay 0·77% (0·0198/2·57 × 100) more rent per square foot to have one more neighbour in the first ring. Or, firms would be willing to pay just under half a S.D. increase in unit rent (or 0·077 of 0·17) for one S.D. increase in immediate neighbours. If a firm moves from the fringe with no neighbours to the height of the action in our sample with 50 immediate neighbours, it would be willing to pay over 2 S.D.s in unit rent increases. Of course, that move could bring more

13. 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).

14. Note the S.D. and average of ring counts rise as we move out (S.D.s being 20 and 49 for rings 2 and 3). This 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 metre increment) but (r_1+r_2) rises with distance. Table B1 in Appendix B reports the agency count and its S.D. in our sample for the rings.

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

Finally, the remaining variable is the number of commercial establishments of all types in a tract. This could represent a local agglomeration economy, or a market potential and shopping centre effect. We will show below that such effects do not extend beyond the first ring. The variable could also represent the potential number of slots available for a new agency, where more slots offer more location choices (more drawings of ε_{jz} from which to choose the highest value within the neighbourhood). In a nested logit context, the coefficient implies a significant positive correlation between error drawings within tracts. The effect of neighbourhood scale is large: a 1% increase in the number of commercial establishments in the tract raises expected births by 0·5%. From a mean of 400 establishments per tract, a firm would be willing to pay a 0·05% increase in price per square foot of office space for one more commercial establishment.

3.3. *Identification*

The instruments we discussed earlier are specifically listed in the footnote to Table 4. These instruments include 1970 housing market variables and counts of advertising agencies for different rings in 1977. All 1992 covariates are treated as endogenous except for own ring total commercial 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 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 terms average 0·50.

The chief concern is the validity of using particular 1977 ring stock variables as instruments, especially for the first ring. The issue is that we have asserted some persistence in location patterns of advertising agencies over time due to moving costs so that 1977 stocks are correlated with 1992 ones; we have also asserted that tract unobservables change over time so 1992 errors are uncorrelated with 1977 stocks, a seemingly delicate balancing act. Thus we explore the validity of our assumptions in some detail. First we note that, not only can Sargan tests not reject validity of instruments, but Sargan values in columns (2) and (4) of the table are extremely low, consistent with the specification and our assumption that current errors are orthogonal to past stocks. Sargan tests are subject to false positives; at worst our strategy ameliorates the problem of contemporaneous missing variables. However, we believe the strategy is valid, with our belief reinforced by other experiments detailed in Arzaghi and Henderson (2006) and briefly summarized here.

As an instructive, informal test, first we added the 1977 advertising agency stock counts for the five rings, which are used as instruments in the IV estimation, to the ordinary Poisson in column 3 of Table 4 as covariates. We compare the coefficients on the 1992 ring stock covariates, with and without the inclusion of 1977 ring stock variables. If 1977 ring stock variables are not valid instruments (are correlated with error terms), the coefficients of 1992 stocks should decline. That is, the instruments as covariates will absorb part of the upward bias in the 1992 coefficients if they are correlated with unobserved characteristics of tracts in 1992 (which bias the 1992 coefficients in the first place). For the key 1992 stock variables (*i.e.* for the first two rings), there is no decline in coefficients (0·023 and 0·017 with 1977 variables added in, vs. 0·021 and 0·017 without, for rings 1 and 2 respectively); moreover, the 1977 stocks have zero coefficients (0·00046 and 0·0035 for rings 1 and 2). In a similar type of informal test, in the formulations in columns 3 and 4 of Table 4, we divided 1992 stocks into new (post-1982 birth date) vs. 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 terms. In IV estimation with so many variables, the results are not statistically strong; but the

TABLE 5

Alternative instruments and births vs. deaths

	Births 1992–1997 IV-GMM (1)	Births 1992–1997 CA instrument for the first ing IV-GMM (2)	Deaths 1992–1997 Poisson (3)	Deaths 1992–1997 IV-GMM (4)
Ln(total no. establishments)	0·498**	0·530**	0·366**	0·0600
	(0·117)	(0·110)	(0·101)	(0·209)
Ln(rent/sq. ft.)	−2·57**	−1·98**	0·164	−0·0121
	(0·773)	(0·578)	(0·364)	(0·819)
Stock of agencies 0–250 m.	0·0198**	0·0167**	0·0598**	0·0756**
	(0·00758)	(0·00626)	(0·00578)	(0·0140)
[Willingness-to-pay/sq. ft. %]	[0·77]	[0·84]		
Stock of agencies 250–500 m.	0·0228**	0·0169**	0·00822**	0·00864
	(0·00538)	(0·00304)	(0·00335)	(0·00617)
Stock of agencies 500–750 m.	0·00419	0·00401**	−0·000208	−0·00182
	(0·00259)	(0·00147)	(0·00,176)	(0·00273)
Stock of agencies 750–1000 m.	−0·00149	−0·00212	0·00447**	0·0111**
	(0·00274)	(0·00,181)	(0·00188)	(0·00352)
Stock of agencies 1000–1250 m.	−0·00241	−0·000949	−0·00373**	−0·00182
	(0·00251)	(0·00167)	(0·00183)	(0·00199)
N	164	164	164	164
Pseudo R^2	—	—	0·549	—
Sargan (p-value)	1·71 (0·945)	6·11 (0·411)	—	13·0 (0·042)
First-stage average R^2	0·50	0·36	—	0·50
Average F (min F)	13·0 (7·9)	11·3 (6·1)	—	13·0 (7·9)

Notes: 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 five 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. 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 ring variables are assumed exogenous in column 2. **Significant at 5% level; *significant at 10% level.

pattern of dominance by new stocks disappears as it should because IV estimation removes any correlation of new stocks with current shocks.

Next in Table 5, we explore the validity of instruments and our approach further, considering alternative instruments and whether births replicate deaths. In this table, column 1 repeats the base case from Table 4, column 4.

Alternative instruments. 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. This leads to a modestly weaker first stage F-statistic (6·1) for the own ring stock (vs. 7·9); and the Sargan p-value remains satisfactory. 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 pattern of coefficients and the willingness-to-pay additional rents for additional neighbours are similar to the base case (higher for the own ring).

Deaths. A concern with using the 1977 own ring stock variable as an instrument as noted above is that births may simply replace deaths and location patterns replicate over time, driven by unchanging economic conditions. Here we examine deaths, estimating an equation for agency deaths from 1992–1997 using the same set of covariates. Deaths are agencies that disappear entirely (not movers within Manhattan). In data where replication is a problem, death equations have the same coefficients as birth equations, since the location patterns of births and deaths closely correspond. Results for our context are reported in columns 3 and 4 of Table 5, where, since this is a new model, we present both ordinary and IV results. To start with, the equation fails the Sargan test, which as a joint test on instruments and specification, indicates the death model is poorly specified. This is not surprising. From Davis, Haltiwanger and Schuh's (1996) work, we know death and birth location patterns are generated by different processes. Individual deaths occur mostly for idiosyncratic reasons, in particular special circumstances involving the specific entrepreneur. As such, agencies' deaths by tract should be roughly proportional to tract stocks, so the first ring coefficient should be high, which it is at 0·076. The coefficient is also consistent with the notion that about half of stocks die in each time interval: if tract stocks rise from one to mean stock per tract of about six, deaths increase in the tract by about 45%. The ordinary Poisson (column 3) and IV estimates (column 4) show two things of interest. First, the death specification definitely does not mimic the birth one: location patterns are not simply replicating. In column 4, the effects of rent, total establishments, ring 2, and ring 3 disappear completely. Second, the specification suggests deaths are largely idiosyncratic. For example, deaths do not increase as rent costs rise nor decline as neighbourhood scale rises, as might be expected.

4. ROBUSTNESS OF RESULTS

In Tables 6 and 7, we explore aspects of robustness of our results. We focus on two sets of issues. First, what is the effect of controlling for other types of tract and ring level activities? Second, how sensitive are results to the precise specification of scale effects? We also look at the effect of excluding, from the choice set, tracts in Manhattan that have no history of advertising agencies. In the last subsection, we comment on a variety of other relevant issues.

4.1. *Neighbourhood conditions*

Table 6 starts by exploring the effect on own ring stock and rent coefficients of adding measures of proximity in Manhattan to various activities other than SU advertising agencies. Column 1 of Table 6 repeats our base case estimates from column 4 of Table 4.

4.1.1. Control for total commercial activity. Column 2 of Table 6 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. 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 the own industry or greater choice among own tract commercial rental properties in the nested logit approach, as discussed in Section 2.2. However, column 3 shows that having more commercial establishments in rings beyond the own ring generates no benefits, consistent with an underlying nested logit specification. If total commercial establishments measure market potential externalities, it would be odd to have such benefits dissipate entirely within 250 metres.

4.1.2. Related economic activities. We explore a variety of more specific aspects of access to other commercial activities: activities that represent clients (headquarters), inputs (graphic services), and outlets (media). For these we had difficulty finding any significant effects, and so we widened the net and strengthened effects by combining the first two rings and defining possible effects for these other activities to occur in 0–500 metres. Of the possibilities, only access to broadcasters matters, which represents the new economic geography costs of trade relevant within Manhattan (negotiating and buying advertising time).

TABLE 6

Robustness

	Base case IV-GMM (1)	No control for total estab. IV-GMM (2)	Total estab. in other rings IV-GMM (3)	Broadcasting employ. IV-GMM (4)	Madison Avenue IV-GMM (5)	"Relevant" tracts IV-GMM (6)
Ln(total no. estab.)	0·498**	—	—	0·563**	0·498**	0·458**
	(0·117)			(0·120)	(0·126)	(0·110)
Ln(rent/sq. ft.)	−2·57**	−3·60**	−2·34**	−2·60**	−2·74**	−1·96**
	(0·773)	(1·07)	(1·22)	(0·850)	(1·00)	(0·719)
Stock of agencies 0–250 m.	0·0198**	0·0321**	0·0214**	0·0221**	0·0188**	0·0176**
	(0·00758)	(0·00943)	(1·22)	(0·00916)	(0·00898)	(0·00747)
[Willingness-to-pay/sq. ft. %]	[0·77]	[0·99]	[0·91]	[0·85]	[0·69]	[0·90]
Stock of agencies 250–500 m.	0·0228**	0·0341**	0·0213**	0·0190**	0·0236**	0·0178**
	(0·00538)	(0·00631)	(0·00738)	(0·00530)	(0·00718)	(0·00524)
Stock of agencies 500–750 m.	0·00419	0·00874**	0·000769	0·00281	0·00426	0·00399*
	(0·00259)	(0·00322)	(0·00604)	(0·00261)	(0·00288)	(0·00224)
Stock of agencies 750–1000 m.	−0·00149	0·00450	—	−0·00298	−0·00122	−0·00270
	(0·00274)	(0·00416)		(0·00292)	(0·00315)	(0·00254)
Stock of agencies 1000–1250 m.	−0·00241	−0·00484	—	−0·000954	−0·00248	−0·00146
	(0·00251)	(0·00331)		(0·00290)	(0·00276)	(0·00224)
Ln (total estab.) 0–250 m.			0·379**			
			(0·178)			
Ln (total estab.) 250–500 m.			−0·0529			
			(0·0845)			
Ln (total estab.) 500–750 m.			0·251			
			(0·370)			
Broadcast emp. 0–500 m.				0·000103**		
				(0·000035)		
Madison Avenue dummy					0·0558	
					(0·339)	
N	164	164	164	164	164	138
Sargan (p-value)	1·71 (0·945)	6·03 (0·537)	3·29 (0·655)	1·57 (0·954)	1·36 (0·968)	1·84 (0·934)
First-stage average R^2	0·50	0·60	0·33	0·55	0·44	0·48
Average F (min F)	13·0 (7·9)	20·0 (13·2)	8·3 (6·1)	16·3 (7·4)	10·3 (5·9)	10·1 (5·4)

Note: 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 five 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. For columns 1, 2, 5, and 6, instruments are as in Table 4. Column 4 adds 1977 stocks of headquarters, broadcasters, and graphics to the list of instruments. no., number; estab., establishments; emp., employment; **significant at 5% level; *significant at 10% level.

In column 4 of Table 6, we measure broadcasting scale by total employment in broadcasting; this works better than using a count of establishments given the vast dispersion in broadcast establishment sizes. The IV results show that a one S.D. increase in broadcasting employment (1785) raises expected births by 18%, a noticeable effect (although the ratio of S.D. 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.

For other controls, for buyers, we noted earlier that, by 1992, access to buyers, or headquarters, is unrelated to location choices within Manhattan. Most advertising agency output in terms of design of advertising campaigns is now exported, given changes in costs and technology of communication and travel. Thus the coefficient (not reported in Table 6) on the count of headquarters within 500 metres is negative and insignificant in both ordinary Poisson and IV estimation. For graphic services inputs, in IV estimation, the partial F in first stage regressions for this variable is a dismal 2·26. Thus we considered only ordinary Poisson results. As reported in Arzaghi and Henderson (2006), there is no effect on ordinary Poisson results from including graphic services nearby and it has a coefficient of 0. Basically, graphic services are ubiquitous in Manhattan. We also looked at establishments of MU advertising agency firms and their effect on birth patterns of SU firms. Ordinary Poisson results suggest that having MU advertising agencies nearby helps in the first ring (significant coefficient of 0·027) and hurts in the second; but IV results give small and insignificant coefficients (first ring coefficient of 0·0045).

In summary, in Table 6, we find no evidence that inclusion of other suspected relevant local activities diminishes the estimated benefits of having more advertising agencies in successive rings.

Address signalling effects. In column 5 of Table 6 we looked for address signalling effects. As new firms just starting out, we do not expect address effects for births, in contrast perhaps to huge established firms that do business with America's largest corporations. In column 5, we add a dummy for just the five tracts below Central Park, which include Madison Avenue. These tracts contain the blocks along Madison Avenue that have traditionally been the prestige addresses for advertising agencies. The basic IV results are little affected and the coefficient on the Madison Avenue dummy, while positive, is small and insignificant. If we narrow the dummy to the three most prestigious of these tracts, the coefficient becomes negative (and insignificant). We also experimented with distance measures, such as distance to Rockefeller Center, and found no significant effects.

Relevant tracts. Finally in column 6 of Table 6, we restrict the sample of tracts to 138 tracts that have had at least one advertising agencies present at some point from 1977 on. 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,[15] the willingness to pay for neighbours rises modestly.

4.2. *Alternative specifications of scale effects*

We have assumed scale effects are linear and homogeneous and are represented by counts of advertising agencies. We explore these assumptions next.

4.2.1. Linearity of effects. We experimented with quadratic and log-linear formulations of effects of stock ring counts on profits. We have not highlighted these results because their IV versions are statistically weak.[16] But for ordinary Poissons, both formulations have suggestive results. Under log-linearity, the coefficients (and S.E.s) for the first three rings are 0·391 (0·066), 0·205 (0·070), and 0·140 (0·0603), which 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 0·0470 and −0·000509 (both significant at 10% or better level). If we increase the number of neighbours by a S.D. from 6 to 16, expected births rise by 41% in the first ring. Positive marginal effects persist up to 45 neighbours, which interestingly coincides with the highest concentration of agencies in Manhattan (*i.e.* 51 agencies in the first ring).

15. In general, the coefficients in discrete choice models—the generalized exponential model in our case—are scaled by the S.D. of the residuals. As we exclude the tracts with no births, we expect that the S.D. of the residuals and therefore the magnitude of coefficients decreases.

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

TABLE 7

Robustness: scale measures

	Births 1992–1997 IV-GMM (1)		Births 1992–1997 median emp. Poisson (2)	Births 1992–1997 average emp. Poisson (3)	Births 92–97 Poisson (4)
Ln(total no.	0·248**		0·378**	0·339**	0·427**
establishments)	(0·0987)		(0·0746)	(0·0740)	(0·0700)
Ln(rent/sq. ft.)	−1·07**		−0·801**	−0·912**	−0·887**
	(0·432)		(0·288)	(0·280)	(0·283)
Ln(tot. agency emp.)	0·238**	Ln(agency emp.)	0·0659	0·153**	
0–250 m.	(0·104)	0–250 m.	(0·00423)	(0·0570)	
Ln(tot. agency emp.)	0·352*	Ln(agency emp.)	0·0292	0·112*	
250–500 m.	(0·215)	250–500 m.	(0·0881)	(0·0632)	
Ln(tot. agency emp.)	0·190*	Ln(agency emp.)	0·148	0·0895	
500–750 m.	(0·107)	500–750 m.	(0·109)	(0·0713)	
Ln(tot. agency emp.)	−0·291**	—	—	—	
750–1000 m.	(0·118)				
Ln(tot. agency emp.)	0·0945	—	—	—	
1000–1250 m.	(0·144)				
		Stock agencies	0·0215**	0·0178**	0·0214**
		0–250 m.	(0·00423)	(0·00452)	(0·00418)
		Stock agencies	0·0142**	0·00994**	0·0149**
		250–500 m.	(0·00250)	(0·00283)	(0·00247)
		Stock agencies	0·00261**	0·00380**	0·00265**
		500–750 m.	(0·00112)	(0·00121)	(0·00112)
N	164		164	164	164
Pseudo R^2	—		0·50	0·51	0·500

Notes: 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 five 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. no., number; emp., employment; **significant at 5% level; *significant at 10% level.

4.2.2. The source of scale effects. A critical issue is measuring spillover effects with just the count of enterprises. The reasoning 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 can also serve double duty, allowing for the possibility that, even if establishment counts accurately record the opportunities for networking, quality of neighbourhood establishments may matter. If establishment size is a measure of quality, this decomposition allows us to test for the existence of either employment sources of externalities or quality of neighbour effects.

The basic problem in proceeding with 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 the 1977 variables on average agency size by ring produces first-stage regressions for average establishment size with F-statistics under 2·0. Similarly, if we change the spillover covariates to total advertising agency 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 column 1 of Table 7, 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. The results show positive employment effects for the first three rings now in elasticity form.

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 vs. 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 and the last column in Table 7 defines the base case where we control for just ring establishment counts. Adding median employment by ring in column 2 of Table 7 shows no effect on establishment count coefficients relative to the last column, and yields insignificant coefficients for the employment size measures. However, adding average size measures does reduce the establishment count coefficients for rings 1 and 2 by 18% and 30% respectively in column 3 compared to the last column (with little effect on the rent coefficient). Average employment size variables are significant for ring 1 and almost so for ring 2. Thus, there is at least a hint that employment sizes as well as establishment counts matter. We cannot sort this out for IV estimation, and conceptually, we prefer the establishment count formulation.

4.2.3. Heterogeneity. The idea that firms, in choosing locations within a city, are trading-off rent costs against the benefits of being "at the centre of action" in a large cluster relative to operating on the fringes raises the issue of heterogeneity. Which firms choose to pay higher rents in order to have better neighbourhood networking—are these higher- or lower-quality firms?[17] We explored this issue in Arzaghi and Henderson (2006). A measure of quality is size, where higher-quality firms with, say better match possibilities, have more sales and employment.[18] We cannot introduce heterogeneity for births, because we have no *ex ante* information on firms (and the firms themselves may not have learned their quality). *Ex post* quality measures such as size are likely to be influenced by contemporaneous error drawings. However, we have a sample of 82 movers who change location between 1992 and 1997 within Manhattan. These experienced firms have informed notions of their own quality for which we have a size measure: payroll in their prior location in 1992 (using firm employment yields similar results). That variable is arguably uncorrelated with the error drawings in the new location.

17. Heterogeneity raises the possibility of a separating equilibrium where high-quality firms segment themselves away from low-quality firms by locating in tracts that are too expensive for low-quality firms (Arzaghi, 2005). Segmentation allows high-quality firms to avoid costly investments in potential networking with inferior partners. This notion may be part of the answer to a question: if location choices of advertising agencies are driven by the desire to network with other agencies, why not cluster more in low-rent areas on the fringes of Manhattan, instead of clustering in higher rent areas? We have the logit explanation of firm-tract specific tastes, and Tables 4 and 6 reveal that advertising agencies also value being in tracts with high concentrations of overall commercial activity and being near broadcasters. But the notion of separating equilibria adds another consideration.

18. Differences in firm quality could be reflected in the specification of matching parameters. Assume each firm is endowed with its own "communication ability" parameter u_j. 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

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

where $G(u)$ is the distribution function. Higher-quality firms have higher $E[(u)^{1/(1-\gamma)}]$; they benefit more from neighbours in equations, so their scale effects are enhanced. They also benefit more from being in a neighbourhood with better-quality neighbours, where the distribution function is improved in the sense of first-degree stochastic dominance. Higher-quality firms will have more profits and be larger (more volume of communications) and thus in our interpretation will have larger employment.

To study firm heterogeneity, we follow the approach in Berry (1994) and Bayer, McMillan and Rueben (2004) and postulate a heterogeneous coefficients model, employing a two stage IV estimation procedure. We explore whether higher quality firms have different coefficients for different covariates. Specifically, in the profit function appearing in (4), we assume the coefficient vector ϕ for neighbourhood covariates now takes the form $\phi^k = \phi_0 + \phi_1 q^k$. ϕ^k are firm k's specific coefficients on the X_j covariates, where those coefficients vary with firm quality, q^k, as measured by firm payroll in the prior location. The sample size is small and covers only 40 tracts in terms of location choices of movers. Two findings emerge nonetheless. First coefficients do vary by firm quality. In particular, higher quality firms are willing to pay substantially more in rents for more neighbours than lower-quality firms. The coefficient on the own ring stock variable rises with quality ($\tilde{\phi}_1$ = 0·0175 and $\tilde{\phi}_0$= −0·0490, both significant, where the mean value of q is 3·3), while the absolute value of the rent coefficient declines ($\hat{\phi}_1$ = 1·05 and $\hat{\phi}_0 = -5{\cdot}35$, both significant). Second, if we evaluate overall effects (coefficients of the X_j) at mean quality of movers, magnitudes are similar when we estimate the model assuming homogeneity of effects.

4.3. *Other considerations*

There are a variety of other aspects of estimation and formulation of networking benefits explored in Arzaghi and Henderson (2006). First, for a very small sample of 30 existing (non-birth) firms in 1992, we estimated the effects of covariates on *ex post* absolute profits, getting results consistent with the births one, albeit for a tiny sample.[19] Second, we have 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, we also explored alternative specifications (such as rings having equal areas) to show our results are robust.

Third, when using data at this fine level of spatial detail and 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 all of them. 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 this problem. Even if the instruments are historically spatially correlated, the issue is that shocks today should be independent of these historical variables. Nevertheless, we checked for spatial correlation, using a Moran I test on the residuals from both the ordinary Poisson and the IV-GMM models in Table 4, columns 3 and 4. The Moran I test examines whether a regression of neighbours' average residuals on own tract residuals is significant. Near neighbours in the averaging are given a weight of 1 and others a weight of 0. In denoting what is a near neighbour, the usual procedure is to think of each tract as having eight neighbours on a lattice. To implement this idea, we use a radius from the own tract that gives an average of 8·3 tracts as neighbours. For this radius, the hypothesis of no correlation cannot be rejected, with a p-value of 0·10 for an ordinary Poisson and 0·11 for IV-GMM. Thus, while there may be within tract correlation, the problem of cross tract correlation of error terms does not seem compelling.

Finally, we examined results for MU advertising agencies. These are reported in Table B2. MUs appear to enjoy networking benefits with other MUs for at least the first and possibly the second ring. However, estimation indicated that they do not benefit from networking with SUs.

19. Profit function estimation gives the absolute magnitude of profit function parameters. Results in Arzaghi and Henderson (2006) indicate that the discrete birth and continuous profit function models yield consistent results on relative magnitudes of profit function coefficients. The profit function results also have larger absolute values of coefficients on rents and own ring stock than in the discrete choice formulation.

5. CONCLUSIONS

The fundamental findings in the paper are that, for advertising agencies, (1) at a micro-spatial level, scale externalities are very large but they dissipate very quickly with distance and are gone by 750 metres and (2) benefits of being in a better neighbourhood are capitalized into land rents. The results strongly suggest the importance of networking or information spillover effects in high-end service industries, such as the advertising agencies, which are strongly concentrated in the largest cities. Such benefits may not apply to other industries such as standardized manufacturing found in smaller cities.

The results raise two issues. First, the 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. This issue is especially critical for industries found in the largest metropolitan areas. 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. For advertising agencies, they could represent labour market or shopping-centre externalities at the metropolitan area level; but they might also represent greater opportunities for clustering and networking in larger metropolitan areas. A research challenge for future work is how to sort out the spatial scales at which different types of agglomeration effects operate.

APPENDIX A: OFFICE RENT INTERPOLATION FOR NEW YORK CITY

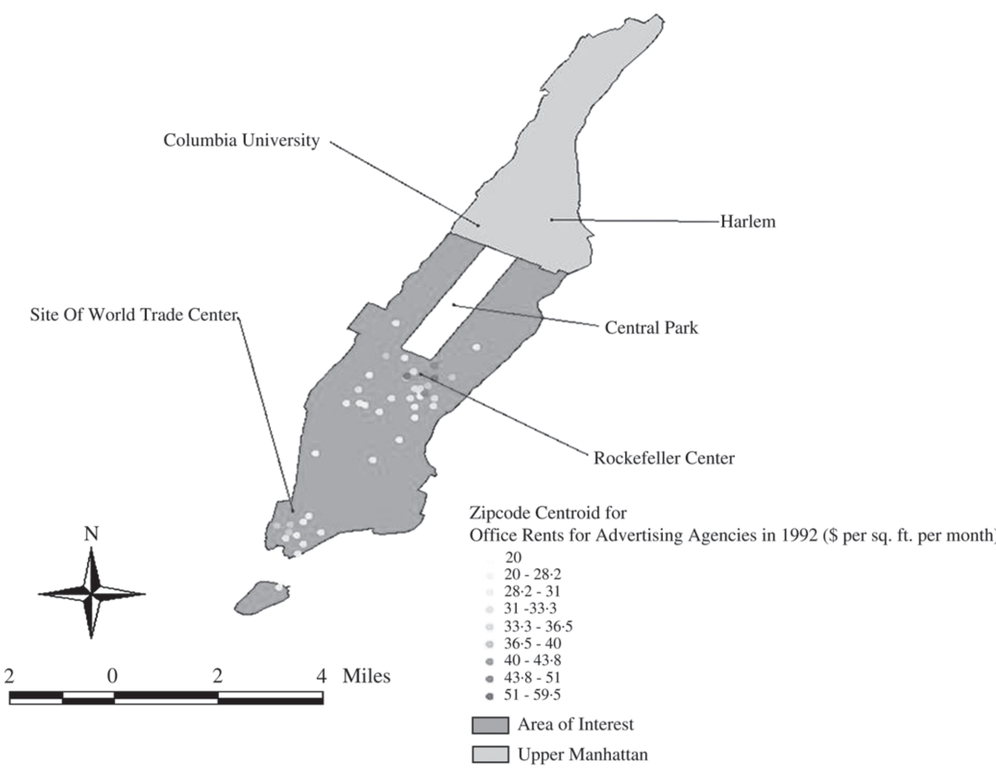

FIGURE A1

Class A office rents for lower Manhattan ZIP codes

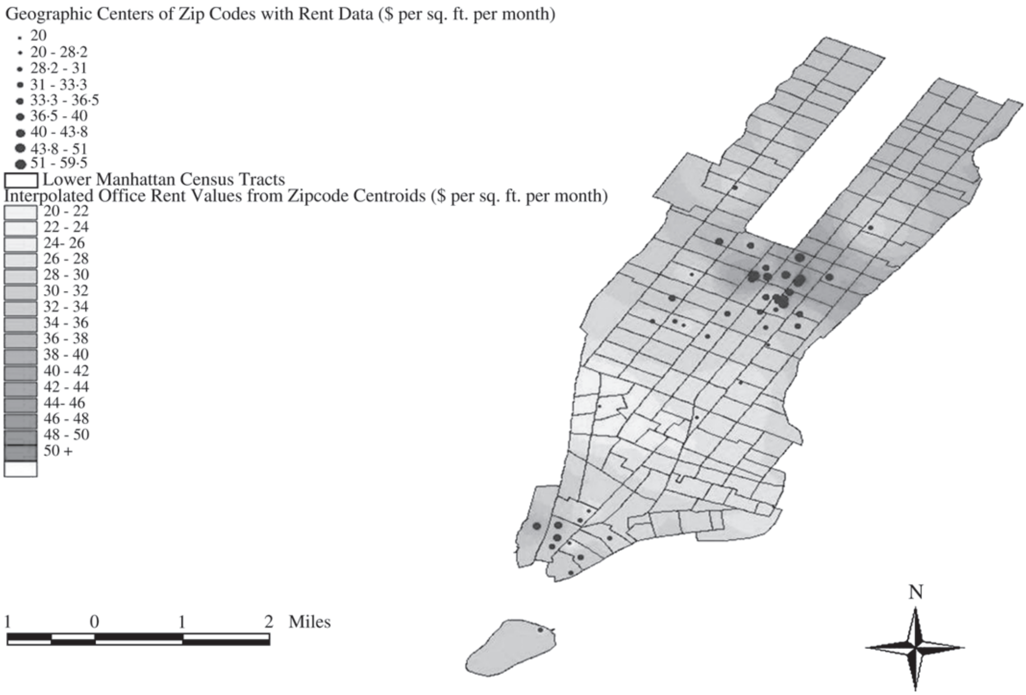

FIGURE A2

Class A office rent surface for lower Manhattan

We use the rent for class A office space as an appropriate rent measure for advertising agencies. The data were provided to us by Torto Wheaton Research. It covers 51 ZIP codes in Manhattan in 1992. The data include the asking 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 construct 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 interpolate the rent for a tract by averaging the value of the rent surface over the tract.

APPENDIX B: BASIC STATISTICS FOR COVARIATES AND OTHER RESULTS

TABLE B1

Means and S.D.s of tract attributes

Tract attribute	Raw mean	Raw S.D.	Weighted mean	Weighted S.D.
Stock of SU agencies 0–250 m.	5·80	9·74	17·5	15·8
Stock 250–500 m.	17·3	20·3	36·8	22·2
Stock 500–750 m.	37·4	49·0	91·4	69·0
Stock 750–1000 m.	34·3	34·6	52·6	33·4
Stock 1000–1250 m	54·3	52·9	96·3	61·4
Ln(total no. establishments)	6·07	1·3	7·1	0·88
Ln(rent/sq. ft.)	3·32	0·17	3·36	0·19
Ln(median payroll of SU agencies)	3·30	2·73	5·24	1·98
N	164	164	502	502

Notes: Attributes are weighted using the number of births in tracts. SU, single unit.

TABLE B2

Results for births of establishments of multi-unit (MU) firms

	Ordinary Poisson	IV-GMM
Ln(total no. establishments)	0·528**	0·679
	(0·230)	(0·574)
Ln(rent/sq. ft.)	−0·1·61**	−2·68
	(0·807)	(2·98)
Stock MUs < 250 m.	0·262**	0·343**
	(0·0439)	(0·111)
Stock MUs 250–500 m.	0·118**	0·0865
	(0·0248)	(0·0611)
Stock MUs 500–750 m.	0·0216	−0·00216
	(0·0190)	(0·0498)
Stock MUs 750–1000 m.	−0·0461*	0·00271
	(0·0257)	(0·0696)
Stock MUs 1000–1250 m.	−0·0215	−0·00187
	(0·0197)	(0·0406)
Constant	0·0228	2·00
	(3·05)	(12·0)
N (births)	164	164
Pseudo R^2	0·460	
Sargan p-value		0·206

Notes: In the table, we examine the results for MU advertising agencies, following the formulation for single units (SUs) in Table 4, except now we specify MUs to benefit from networking with other MUs. Column 1 contains the ordinary Poisson and column 2 the IV for baseline specifications. There are strong scale effects in rings 1 and 2 for ordinary Poissons and in ring 1 for IV results. In the basic formulations, we tried adding 1, or 3 and 5 SU ring variables. In the ordinary Poisson and IV-GMM estimations, we never obtained positive, even modestly significant results for SU variables. Their inclusion has no impact on the coefficients of baseline covariates. no., number; MU, multi-unit; **significant at 5% level; *significant at 10% level.

Acknowledgements. 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, Leo Feler, and William Strange on aspects of the paper and from participants in seminars at Brown University, LSE, NBER Real Estate and Local Public Finance Meetings, Yale University, University of Minnesota and University of Toronto. We thank the referees and the editor, Andrea Prat, for helpful comments in preparing a final version. We thank NSF for financial support of this project (award number 0111803) and Arzaghi thanks the Kennedy School of Government at Harvard University and the Dubai School of Government for a research fellowship and financial support during the final revisions of this paper.

REFERENCES

ARZAGHI, M. (2005), "A Study of Networking and Location Decisions in the Advertising Agency Industry" (PhD thesis, Brown University).

ARZAGHI, M. and HENDERSON, J. V. (2006), "Networking Off Madison Avenue" (http://www.econ.brown.edu/faculty/henderson/papers/MadisonAve101206.pdf)

BALA, V. and GOYAL, S. (2000), "A Noncooperative Model of Network Formation", *Econometrica*, **68**, 1181–1229.

BAYER, P., McMILLAN, R. and RUEBEN, K. (2004), "Residential Segregation in Equilibrium" (Mimeo, Yale University).

BERNDT, E. R. and SILK, A. J. (1994), "Costs, Institutional Mobility Barriers, and Market Structure: Advertising Firms as Multiproduct Firms", *Journal of Economics and Management Strategy*, **3**, 437–450.

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

CONOVER, W. J. (1999), *Practical Nonparametric Statistics*, 3rd edn (New York: John Wiley & Sons).

DAVIS, S., HALTIWANGER, J. and SCHUH, S. (1996), *Job Creation and Destruction* (Cambridge: MIT Press).

DUDEY, M. (1990), "Competition by Choice: The Effect of Consumer Search on Firm Location Decisions", *American Economic Review*, **80** (5), 1092–1104.

FUJITA, M., KRUGMAN, P. and VENABLES, A. J. (1999), *The Spatial Economy: Cities, Regions and International Trade* (Cambridge: MIT Press).

GLAESER, E. and GYOURKO, J. (2002), "The Impact of Zoning on Housing Affordability" (Harvard Institute of Economic Research, Paper No. 1948, Harvard University).

GLAESER, E. and MARE, D. (2001), "Cities and Skills", *Journal of Labor Economics*, **19**, 316–342.

GOYAL, S. and MORAGA-GONZALEZ, J. L. (2001), "R&D Networks", *RAND Journal of Economics*, **32** (4), 686–707.

GUIMARÃES, P., FIGUEIRDO, O. and WOODWARD, D. (2003), "A Tractable Approach to the Firm Location Decision", *Review of Economics and Statistics*, **85**, 201–204.

HAMEROFF, E. (1998), *The Advertising Agency Business* (New York, NY: McGraw-Hill).

HENDERSON, J. V. (2003), "Marshall's Scale Externalities", *Journal of Urban Economics*, **53** (1), 1–28.

HOLMES, T. (2005), "The Location of Sales Offices and the Attraction of Cities", *Journal of Political Economy*, **113** (3), 551–581.

JACKSON, M. and WOLINSKY, A. (1996), "A Strategic Model of Social and Economic Networks", *Journal of Economic Theory*, **71**, 44–74.

JAFFE, A. B., TRAJTENBERG, M. and HENDERSON, R. (1993), "Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations", *Quarterly Journal of Economics*, **108**, 577–598.

KOLKO, J. (2000), "Essays on Information Technology, Cities, and Location Choice" (PhD Thesis, Harvard University).

LUCAS, R. E., Jr. and ROSSI-HANSBERG, E. (2002), "On the Internal Structure of Cities", *Econometrica*, **70**, 1445–1476.

MARSHALL, A. (1890), *Principles of Economics* (London: Macmillan and Co.).

McMILLEN, D. P. and SINGELL, L. D., Jr. (1992), "Work Location, Residence Location, and the Intraurban Wage Gradient", *Journal of Urban Economics*, **32**, 195–213.

MULLAHY, J. (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.

READ, T. R. C. and CRESSIE, N. A. C. (1988), *Goodness-of-Fit Statistics for Discrete Multivariate Data* (New York: Springer-Verlag New York Inc.).

ROSENTHAL, S. and STRANGE, W. (2004), "Evidence on the Nature and Sources of Agglomeration Economies", in J. V. Henderson and J.-F. Thisse (eds.) *Handbook of Regional and Urban Economics* (Amsterdam: North Holland) 2119–2171.

ROSENTHAL, S. and STRANGE, W. (2006), "The Attenuation of Human Capital Spillovers" (Working paper; http://www.rotman.utoronto.ca/∼wstrange/Rosenthal_Strange_Wage_11-29-06.pdf), University of Toronto.

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

TRAIN, K. E. (2000), *Discrete Choice Methods with Simulation* (Cambridge, UK: Cambridge University Press).

WINDMEIJER, F. and SILVA, J. S. (1997), "Endogeneity in Court Data Models: An Application to Health Care", *Journal of Applied Econometrics*, **12**, 281–294.