Imitation in location choice

Berg, Nathan

2008

Imitation in location choice

Nathan Berg*

Abstract: Under the assumption of perfect competition, it is difficult to avoid the conclusion that abandoned properties and long undeveloped neighborhoods remain that way because they are unprofitable. In contrast, this paper introduces a model in which firms systematically overlook neighborhoods with little commercial activity because of a positive informational externality motivating later movers to condition choice of location on earlier movers' locations. When this occurs, firms sometimes find it profitable to imitate early movers' locations even though privately acquired information suggests locating elsewhere. The model facilitates normative analysis of imitation in location choice by explicitly quantifying losses in aggregate efficiency following a shift from centralized to decentralized regimes. The model provides a tool for investigating the hypothesis of inefficient lock-in as it relates to neighborhoods in U.S. urban centers that remain underutilized despite the presence of profitable business prospects.

Imitation is natural to man from childhood, one of his advantages over the lower animals being this, that he is the most imitative creature in the world, and learns at first by imitation. —Aristotle (Poetics, Chapter IV)

1 Introduction

This paper presents a model of information acquisition and location choice among profit-maximizing firms. In contrast to theories based on the assumption of first-mover advantage, the model focuses on the benefit that second movers enjoy by conditioning on first movers' locations. Such opportunities for utilizing information implicit in the locations of other firms give rise to the possibility of imitation, quantified as the extent to which second movers substitute away from costly acquisition of private information into costless observation of first movers.

One motivation for modeling imitation in location choice is to investigate the possibility of behavioral barriers to redevelopment in older, low-income neighborhoods within central cities. Brueckner and Rosenthal (2005) identify age of housing stock as a key predictor of future changes in neighborhood income, which naturally leads to comparisons of factors that influence new development in suburbs versus redevelopment within central cities. Observing low-income neighborhoods in central cities with very little retail activity, the standard neoclassical model suggests that, despite advantages such as lower rents and fewer competitors, these areas are passed over for good reason—because firms cannot profitably operate stores there. In contrast, the hypothesis considered here is that small firms use an imitation heuristic for choosing locations, which succeeds at maximizing profits in environments with abundant information such as thriving big-box developments in the suburbs, but fails to exploit genuinely profitable opportunities in urban environments with little or no available in-

formation concerning revenues and costs. In these low-information environments with limited retail activity, stigmatizing perceptions may result in lock-in, because lack of new entrants into these neighborhoods cuts off the production of new information about the profitability of current business activity.

One of the models results is to reveal conditions under which imitation in location choice is consistent with individual profit maximization. In so doing, the model provides a benchmark for addressing the normative question of how imitation affects aggregate efficiency. Incentives that rationalize imitation from the point of view of individual firms can, however, lead to socially inefficient neglect of profitable locations—for example, when firms do not consider moving into a no-retail neighborhood simply because no other firms are observed there, and not because expected profits were estimated and deemed too low.

For parameterizations with low costs of private information, the models first movers acquire large quantities of private information, and imitation by second movers is consistent with social efficiency (i.e., maximization of aggregate profit by a centrally coordinated program designed to efficiently exploit the positive informational externality flowing from first to second movers). In this case, the first mover is well informed about where to find good locations and, consequently, centralized and decentralized regimes differ very little. As information becomes more expensive, however, imitation becomes increasingly inconsistent with aggregate efficiency, because first movers acquire small quantities of private information and the aggregate benefits of pooling additional private signals grow larger. But because firms and centralized planners both demand very little information when information is very expensive, the gap in aggregate profits between decentralized and centralized regimes shrinks and inefficiency tends toward zero at the other extreme of the cost-of-information spectrum. Thus, efficiency losses are largest in the intermediate range of the cost of private

information and minimal at the extremes.

Interest in theoretical mechanisms that lead to socially inefficient spatial lock-in among retailers draws, in part, on recent evidence that stores such as Starbucks and Home Depot have earned profits far in excess of what their own demand forecast models predicted by investing in long overlooked, low-income neighborhoods previously regarded as unprofitable (Weissbourd, 1999; Helling and Sawicki, 2003; Sabety and Carlson, 2003). For example, Vice President of Starbuck's Store Development Cydnie Horwat writes: "Our Urban Coffee Opportunities joint venture has essentially shown that Starbucks can penetrate demographically diverse neighborhoods in underserved communities, such as our store in Harlem, which is not something that we had previously looked at" (Francica, 2000).

This raises questions. How could Starbucks have overlooked a profitable opportunity for so long, and why did it require a new, joint initiative to discover that the coffee giant could operate profitably in ethnically mixed, low income neighborhoods? Are neighborhoods in central cities that retailers avoid really less profitable, or do interdependencies among firms' location decisions lead to inefficient lock-in at a status quo biased against such neighborhoods simply because firms have decided against them in the past? And finally, should we be surprised that sophisticated firms, even those that conduct extensive market research, base location decisions primarily on observed choices of other firms instead of independently weighing the costs and benefits associated with each of many candidates drawn from a large consideration set?

The model presented below suggests that we should not be surprised to find imitation heuristics in widespread use among business decision makers responsible for choosing locations because, in many environments, imitation is consistent with profit maximization. The model also demonstrates and provides a means of quantifying the social cost of imitation in location choice. The behavioral mechanism of imitation as

a cause for spatial concentration suggests one plausible explanation for why neighborhoods that should be capable of sustaining profitable retail activity—and would do so, if firms conducted independent calculations of expected profit—sometimes wind up in a state of abandonment, failing over a sustained period of time to attract mainstream commercial activity.

It is unclear how far models of crime as a factor in residential location (Helsley and Strange, 1999; Verdier, T., and Zenou, 2004; Helsley and Strange, 2005) extend to the case of retail location choice. Yet beliefs about crime appear to play a large role in conditioning firms' decisions about entering ghettos and other stigmatized neighborhoods (Bray, 2007; Weissbourd, 1999). Interviews with business decision makers responsible for location choice (Berg, 2007) confirm that many firms cite crime as a reason for not considering stigmatized neighborhoods, although those firms rarely, if ever, conduct or commission quantitative benefit-cost assessments to justify such omissions from their consideration sets. Instead, the decision processes of most firms' location decisions appear to rely heavily on imitation heuristics and threshold rules (i.e., satisficing) that quickly narrow down the consideration set to a handful of candidates, in line with the ideas of Simon (1954, 1955), Cyert and March (1963), and March (1988). Indeed, the biology literature shows imitation to be an adaptive strategy for animals in a number of environments (Noble, Todd and Tuci, 2001; Hutchinson, 2005), just as the social science literature identifies environments where imitation leads to success (Gigerenzer and Selten, 2002; Bosch-Domnech and Vriend, 2003) Concerning the normative focus of this paper, it is useful to recall that spatial agglomeration, or clustering, in the classic Hotelling (1929) model is wasteful, as firms locate in the center to split the market rather than at locations minimizing transportation costs. Hotelling, and later Boulding (1996), generalized the idea of socially wasteful agglomerations to a broad range of social settings. This negative

assessment was later tempered by arguments emphasizing the benefits of spatial agglomerations, which help consumers by economizing on shopping costs in terms of shopping time, transportation cost (Eaton and Lipsey, 1976, 1979) and uncertainty reduction (Wolinsky,1983; Dudey, 1990). Following numerous papers on efficiency gains from agglomeration, however, new negative assessments appeared, for example Dudey (1993) on welfare-decreasing agglomerations. More recent theoretical and empirical work emphasizes the role of physical distance in the production function—indivisibility in production (Kanemoto, 1990), labor market pooling (Rosenthal and Strange, 2001), and complementarity between workers and firms (Andersson, Burgess and Lane, 2007)—to better characterize efficiency of agglomerations. Rather than trying to harmonize contradictory normative theories (Fischer and Harrington, 1996), this paper attempts to exploit the mixed normative message in the agglomeration literature. This follows Gigerenzer et al (1999) and Gigerenzer and Selten (2002), whose normative approach—referred to as ecological rationality—seeks to analyze when decision procedures are well matched, or badly matched, to decision environments. The matching concept underlying ecological rationality stands in contrast to universal, context-free and content-blind normative criteria for evaluating rationality, such as transitivity, the Kolmogorov axioms of probability theory, or the Savage axioms of expected utility theory.[1]

---

[1]Context-dependent normative analysis does not imply relativity, as there remain many compelling reasons other than violations of consistency axioms for policy makers to be concerned about behavioral underpinnings of spatial agglomerations. Berry (1961), for example, argues that steepness of city-size distribution curves is inversely related to economic development and, therefore, that policy makers interested in economic development have good reason to be concerned with different forms of agglomeration as a primary issue in planning. Muiz, and Galindo (2005) present evidence on suburban agglomerations and environmental impacts. And Anas and Rhee (2007) demonstrate the sensitivity of normative evaluations of policies that concern spatial agglomerations to apparently innocuous assumptions such as exogenous agricultural land rents in areas surrounding cities. Similarly, Turner (2007) shows that small coordination and free-rider microstructures lead to a large divergence between equilibrium and socially efficient spatial distributions.

The plan of this paper is as follows. Section 2 presents the model and main theoretical results. Section 3 discusses these results in the context of three distinct academic literatures. Section 4 returns to the problem of underutilized resources in urban areas with a conclusion offering interpretations of the model as applied to the case of urban ghettos in the U.S.

2 The Model

The model considers firms that have two primary choice variables: how much private information to acquire about locations, and choice of location. Following extensive theoretical (Prescott and Visscher, 1977; Kogut, 1983) and empirical literatures (Chang, 1995; Chang and Rosenzweig 2001; Chung, 2001) on sequential entry and exit, this model assumes that each firm makes the joint decision of information acquisition and location choice at a single point in a longer sequence comprised of similar joint decisions by other firms. To investigate the effect of a firm's position in this temporal sequence on information acquisition—in particular, the extent to which firms condition their choices of location on previous movers' locations instead of collecting independent information on their own—all firms' objective functions are assumed to differ only in the sets of information available for conditioning expected profit. Thus, heterogeneous values of maximized expected profit arise solely because of heterogeneous positions in the sequence of moves and, consequently, the different sets of information firms acquire.

To fix ideas, the simplest possible temporal sequence of decisions is considered: two firms each of which chooses a quantity of information and location to maximize expected profit. The first mover is referred to as Firm 1. The second mover is referred to as Firm 2, and moves after observing Firm 1's choice of location.

Locations are indexed on the unbounded real line. Because the focus is information acquisition rather than strategic considerations or other interesting problems such as multiple equilibria, Knightian uncertainty, and complex dynamics, the model takes a shortcut by assuming the existence of a unique profit-maximizing location $a \in \Re$, referred to as the *ideal location*. This unique profit-maximizing location, a, is assumed to be the same for both firms.

Information acquisition is important in the model because a is unknown to both firms. Firms must therefore condition predictions of a on private signals they acquire, denoted x_1 and x_2 for Firms 1 and 2, respectively. In addition to x_2, Firm 2 also conditions its expectations of a on the observed location of Firm 1, which is denoted y_1. Firm 2's choice of location is denoted y_2.

Quantities of private information are denoted θ_1 and θ_2, $0 \leq \theta_i \leq 1$, $i = 1, 2$. The "quantity of information" θ_i represents R^2 in a univariate regression of a on the privately acquired signal x_i. Larger θ_i means that Firm i chooses more private information or, equivalently, lower conditional variance of a given x_i.

Privately acquired signals come from a variety of sources, including public data sets and private vendors, both of which incur time, processing and sometimes explicit financial costs. The model captures the costs of acquiring private information with a continuously differentiable and weakly increasing cost function $C(\theta)$, with $C(0) = 0$ and $C'(\theta) \geq 0$. Firm 1's location is a crucial piece of information for Firm 2. Under the assumption that this information is easily observable, it makes sense to keep it distinct from the privately acquired signal x_2 and without any effect on Firm 2's total cost of information $C(\theta_2)$.

If either firm (or both) knew where the best location was (i.e., knew a), profit would be given by the exogenous parameter π_0, interpreted as maximized profit in the ideal case of full knowledge with zero information costs. Given uncertainty about

a, however, firms experience actual costs from two separate categories. First is the cost of deviating from a, which should be interpreted as reduced sales, extra transportation costs, or higher taxes. Second is the cost of private information, which is acquired specifically to reduce the expected deviation of the firms actual location from a. With quadratic costs of deviating from a, the profit function takes the form $\pi_0 - (y_i - a)^2 - C(\theta_i)$, $i = 1, 2$. Because a is uncertain, the *ex ante* objective functions are stated in expected form:

$$\pi_1(y_1, \theta_1) = \pi_0 - \mathrm{E}[(y_1 - a)^2 | x_1; \theta_1] - C(\theta_1), \tag{1}$$

$$\pi_2(y_2, \theta_2) = \pi_0 - \mathrm{E}[(y_2 - a)^2 | y_1(\theta_1), x_2; \theta_2] - C(\theta_2). \tag{2}$$

Note that Firm 1 and Firm 2's expected profit functions differ only in the information upon which expectations are conditioned. The notation makes clear that Firm 1's expectation of expressions involving a is conditioned by its private information x_1, which depends on its choice of θ_1. Firm 2's expectation of expressions involving a is conditioned by Firm 1's location $y_1(\theta_1)$ and Firm 2's privately acquired signal x_2. Firm 2's expectations depend on its choice of θ_2. The notation in equation (2) expresses y_1 as a function of θ_1 to make the dependence of Firm 2's information acquisition on Firm 1's choice of information explicit.

Expected deviations from a, which appear in each firm's profit, can be decomposed as follows:

$$\mathrm{E}[(y_1 - a)^2 | x_1; \theta_1] = (y_1 - \mathrm{E}[a | x_1; \theta_1])^2 + \mathrm{var}(a | x_1; \theta_1), \tag{3}$$

$$\mathrm{E}[(y_2 - a)^2 | y_1(\theta_1), x_2; \theta_2] = (y_2 - \mathrm{E}[a | y_1(\theta_1), x_2; \theta_2])^2 + \mathrm{var}(a | y_1(\theta_1), x_1; \theta_2). \tag{4}$$

Because the first terms on the right hand side of (3) and (4) have unique minima at zero, and because y_i appears nowhere else in Firm i's objective function, optimal location choice rules are given by:

$$y_1^* = \mathrm{E}[a | x_1; \theta_1] \quad \text{and} \quad y_2^* = \mathrm{E}[a | y_1(\theta_1), x_2; \theta_2]. \tag{5}$$

Whenever y_1^* is linear in x_1, then $y_2^* \to y_1^*$ as $\theta_2 \to 0$. The reason for this is that Firm 2 receives an undistorted version of Firm 1's private signal while acquiring no private information of its own, which implies that Firm 2 must have the same expectation of a and consequently choose the same location as Firm 1.

After substituting (3) into (1), the first-order condition for θ_1 is:

$$2(y_1 - \mathrm{E}[a|x_1; \theta_1])\frac{\partial \mathrm{E}[a|x_1; \theta_1]}{\partial \theta_1} - \frac{\partial \mathrm{var}(a|x_1; \theta_1)}{\partial \theta_1} - \frac{\partial C(\theta_1)}{\partial \theta_1} = 0, \qquad (6)$$

but because the global solution for y_1^* in (5) makes the first term in parentheses on the left-hand side of (6) uniformly zero, this first-order condition simplifies to $-\frac{\partial \mathrm{var}(a|x_1; \theta_1)}{\partial \theta_1} = \frac{\partial C(\theta_1)}{\partial \theta_1}$. This requires that Firm 1 choose θ_1 to set the marginal reduction in conditional variance of a equal to the marginal cost of information. A solution to the first-order condition for θ_1 may not always exist and, even when it does, it may not correspond to a global maximum. The boundary values, $\theta_1 = 0$ and $\theta_1 = 1$, must also be checked. The choice $\theta_1 = 0$ maximizes expected profit when information is too expensive to justify its acquisition at any level, in which case Firm 1 locates at the unconditional mean, $y_1 = \mu_a$, and achieves expected profit $\pi_0 - \sigma_a^2$. The choice $\theta_1 = 1$ represents cases where Firm 1 acquires maximally precise information, implying that $\mathrm{var}(a|x_1; \theta_1)|_{\theta_1=1} = 0$, and profit $\pi_0 - C(1)$, which is certain in this case rather than expected.

Given the real-world policy problems associated with the hypothesis of imitation causing inefficient spatial agglomerations that fail to utilize profitable opportunities elsewhere, the most interesting case to consider is when Firm 1 acquires information and Firm 2 does not: $\theta_1 > 0$ and $\theta_2 = 0$, referred to as *absolute imitation.* One seeks to identify the conditions under which imitation of this kind is consistent with expected-profit maximization. Thus, it is assumed that the global maximizer θ_1^* lies on the strict interior of the unit interval: that is, $\pi_1(\mathrm{E}[a|x_1; \theta_1^*], \theta_1^*) > \max\{\pi_0 - \sigma_a^2, \pi_0 - C(1)\}$.[2]

---

[2] A sufficient but not necessary condition for existence of an interior solution for θ_1 is that the conditional variance

Conditions are also sought under which Firm 2 acquires private information, but strictly less information than Firm 1: $0 < \theta_2^* < \theta_1^*$, referred to as partial imitation, because y_2^* and y_1^* are closer on average than they would be if location choices were based solely on private information.

Turning to Firm 2's condition for optimal information acquisition, the first-order condition for θ_2 is:

$$-\frac{\partial \mathrm{var}(a|y_1(\theta_1), x_2; \theta_2)}{\partial \theta_2} = \frac{\partial C(\theta_2)}{\partial \theta_2}, \tag{7}$$

which makes clear the dependence of Firm 2's choice of θ_2 on θ_1. Again, this first-order condition may not have a solution, and its solution may be dominated by choices at the boundaries. In cases where $-\frac{\partial \mathrm{var}(a|y_1(\theta_1), x_2; \theta_2)}{\partial \theta_2} < \frac{\partial C(\theta_2)}{\partial \theta_2}$ for all θ_2, Firm 2's marginal benefit of private information is never greater than its marginal cost. The following condition describes when Firm 2 acquires no private information.

Result 1 (Absolute Imitation): Provided that $C(\theta)$ is convex, that $\mathrm{var}(a|y_1(\theta_1), x_2; \theta_2)$ is convex in θ_2 (i.e., Firm 2's marginal benefit of private information acquisition is decreasing in θ_2), and provided that the following inequality holds:

$$-\frac{\partial \mathrm{var}(a|y_1(\theta_1), x_2; \theta_2)}{\partial \theta_2}\Big|_{\theta_2=0} < \frac{\partial C(0)}{\partial \theta} < -\frac{\partial \mathrm{var}(a|x_1; \theta_1)}{\partial \theta_1}\Big|_{\theta_1=0}, \tag{8}$$

then Firm 2 absolutely imitates Firm 1. Absolute imitation means that Firm 2 acquires no information on its own, $\theta_2^* = 0$, while Firm 1 acquires a strictly positive quantity of information, $\theta_1^* > 0$. If Firm 1's location is linear in its private signal, then this condition also implies that Firm 2 will choose the same location as Firm 1, $y_2^* = y_1^*$.

Condition (8) relies on the fact that, if marginal benefit of private information is less than marginal cost starting from an initial position of zero information, then there is never any incentive to acquire information. Firm 2's marginal benefit is decreasing

of a given x_1 is weakly convex in θ_1, which holds, for example, in case a and x_1 are jointly normal, as shown in later sections.

in own private information acquisition (by convexity of Firm 2's conditional variance), and marginal cost is increasing (by convexity of the cost function). Therefore, the first inequality rules out Firm 2 acquiring private information, and the second inequality ensures that Firm 1 acquires some.[3]

2.1 Joint normality

To examine the interdependence of the firms' levels of information acquisition using tractable functional forms, attention now turns to the consequences of assuming joint normality:

$$a \sim \mathrm{N}(\mu_a, \sigma_a^2), \quad x_1 = a + \epsilon_1, \quad \text{and} \quad x_2 = a + \epsilon_2, \tag{9}$$

where ϵ_i is normal, with mean $-\mu_a$ (so that, without loss of generality, $\mathrm{E}[x_i] = 0$), and independent from a, for $i = 1, 2$. Thus, each firm's acquisition of information can be expressed as:

$$\theta_i \equiv [\mathrm{corr}(x_i, a)]^2 = (\sigma_a^2)^2/(\sigma_a^2 \sigma_i^2) = \sigma_a^2/\sigma_i^2, \tag{10}$$

where $\sigma_i^2 \equiv \mathrm{var}(x_i)$, $i = 1, 2$.

Joint normality allows conditional expectations to be expressed explicitly. Recalling that x_1 has mean zero by definition, Firm 1's conditional expectation of a [and by equation (5), its expected-profit-maximizing location] is:

$$y_1^* = \mu_a + [\mathrm{cov}(x_1, a)/\sigma_1^2]x_1 = \mu_a + \theta_1 x_1. \tag{11}$$

Conditional variance of a given Firm 1's observed signal is given by the convenient formula:

$$\mathrm{var}[a|x_1] = \sigma_a^2 - [\mathrm{cov}(x_1, a)]^2/\sigma_1^2 = \sigma_a^2(1 - \theta_1). \tag{12}$$

The condition under which Firm 1 acquires a positive quantity of information [the second inequality in (8), which requires that marginal cost of the first unit of infor-

---

[3]For more on the value of information and its interactions with risk aversion, not considered further here, see Willinger (1989), Hilton (1981), and Eeckhoudt, Godfroid and Gollier (2001).

mation is less than its marginal benefit] simplifies to $C'(0) < \sigma_a^2$. When this condition is satisfied, the interior solution θ_1^* solves:

$$\sigma_a^2 = C'(\theta_1). \tag{13}$$

After making the substitutions $\sigma_1^2 = \sigma_a^2/\theta_1$, $\sigma_2^2 = \sigma_a^2/\theta_2$, and $\text{var}(y_1) = \theta_1^2(\sigma_a^2/\theta_1) = \theta_1\sigma_a^2$, Firm 2's conditional expectations, $\text{E}[a|y_1, x_2]$ and $\text{var}(a|y_1, x_2)$, can be expressed in terms of θ_1 and θ_2:

$$
\begin{aligned}
\text{E}[a|y_1(\theta_1), x_2; \theta_2] &= \mu_a + [\text{cov}(y_1, a) \ \text{cov}(x_2, a)] \begin{bmatrix} \text{var}(y_1) & \text{cov}(y_1, x_2) \\ \text{cov}(y_1, x_2) & \text{var}(x_2) \end{bmatrix}^{-1} \begin{bmatrix} y_1(\theta_1) - \mu_a \\ x_2 \end{bmatrix} \\
&= \mu_a + [\theta_1\sigma_a^2 \ \sigma_a^2] \begin{bmatrix} \theta_1\sigma_a^2 & \theta_1\sigma_a^2 \\ \theta_1\sigma_a^2 & \sigma_a^2/\theta_2 \end{bmatrix}^{-1} \begin{bmatrix} \theta_1 x_1 \\ x_2 \end{bmatrix} \tag{14} \\
&= \mu_a + \frac{\theta_1(1 - \theta_2)}{1 - \theta_1\theta_2} x_1 + \frac{\theta_2(1 - \theta_1)}{1 - \theta_1\theta_2} x_2, \tag{15}
\end{aligned}
$$

and:

$$
\begin{aligned}
\text{var}(a|y_1(\theta_1), x_2; \theta_2) &= \sigma_a^2 - [\theta_1\sigma_a^2 \ \sigma_a^2] \begin{bmatrix} \theta_1\sigma_a^2 & \theta_1\sigma_a^2 \\ \theta_1\sigma_a^2 & \sigma_a^2/\theta_2 \end{bmatrix}^{-1} \begin{bmatrix} \theta_1\sigma_a^2 \\ \sigma_a^2 \end{bmatrix} \tag{16} \\
&= \sigma_a^2[1 - (\theta_1 + \theta_2 - 2\theta_1\theta_2)/(1 - \theta_1\theta_2)]. \tag{17}
\end{aligned}
$$

Equation (15) implies that the more information Firm 1 acquires, the less weight Firm 2 places on its own private signal (i.e., $\frac{\partial}{\partial\theta_1}[\theta_2(1 - \theta_1)/(1 - \theta_1\theta_2)] = -\theta_2(1 - \theta_2)/(1 - \theta_1\theta_2)^2 \leq 0$). As long as $\theta_1 > \theta_2$, equation (15) also shows that Firm 2 will weight Firm 1's information more than its own, and if $\theta_2 = 0$, locate exactly where Firm 1 does: $y_2^* = y_1^* = \mu_a + \theta_1 x_1$.

The following expression, which is positive, measures the marginal benefit to Firm 2 of its private information:

$$-\frac{\partial\text{var}(a|y_1(\theta_1), x_2; \theta_2)}{\partial\theta_2} = \sigma_a^2(1 - \theta_1)^2/(1 - \theta_1\theta_2)^2. \tag{18}$$

Taking the second derivative with respect to θ_2 reveals a negative sign, which shows convexity of conditional variance [i.e., the second derivative of $\text{var}(a|y_1(\theta_1), x_2; \theta_2)$ without the negative sign is positive], as required in Result 1. To investigate whether the other part of Result 1's absolute imitation condition holds, a specification of the cost function is required, which subsequent sections provide. If both firms choose interior quantities of information, then they will equate marginal benefit of private information with marginal cost. For any strictly increasing cost function, Firm 2 will acquire less private information than Firm 1 if and only if its marginal benefit is less than Firm 1's. This clearly holds in the case of joint normality:

$$-\frac{\partial \text{var}(a|y_1(\theta_1), x_2; \theta_2)}{\partial \theta_2} = \sigma_a^2[(1 - \theta_1)/(1 - \theta_1 \theta_2)]^2 \leq \sigma_a^2 = -\frac{\partial \text{var}(a|x_1; \theta_1)}{\partial \theta_1}. \quad (19)$$

Result 2 (Firm 2 demands less private information than Firm 1): For strictly increasing $C(\theta)$ and joint normality of a, x_1 and x_2, Firm 2 demands less private information than Firm 1: $\theta_2^* \leq \theta_1^*$. This result follows from the fact that Firm 1's marginal benefit of private information is uniformly greater than Firm 2's: $-\frac{\partial \text{var}(a|y_1(\theta_1), x_2; \theta_2)}{\partial \theta_2} \leq -\frac{\partial \text{var}(a|x_1; \theta_1)}{\partial \theta_1}$, as demonstrated in (19).

2.2 Example with exponential cost of information

Suppose the cost function has the following (inverse) exponential form:

$$C(\theta) = -c \log(1 - \theta), \quad c > 0. \quad (20)$$

The key feature of this cost function is that the first unit of information has positive marginal cost c and approaches infinity as θ approaches 1. Solving (13) leads to $\theta_1^* = 1 - c/\sigma_a^2$ for $0 \leq c \leq \sigma_a^2$, and 0 otherwise. Referring back to Result 1, one can observe that, with $c < \sigma_a^2$, the condition for absolute imitation holds:

$$-\frac{\partial \text{var}(a|y_1(\theta_1), x_2; \theta_2)}{\partial \theta_2}\Big|_{\theta_2=0} = \sigma_a^2(1 - \theta_1^*)^2 = (c/\sigma_a^2)c$$

$$< c = \frac{\partial C(0)}{\partial \theta} \leq \sigma_a^2 = -\frac{\partial \text{var}(a|x_1; \theta_1)}{\partial \theta_1}\Big|_{\theta_1=0}. \quad (21)$$

Thus, with the cost function (20), Firm 1 acquires a positive quantity of private information (provided that the marginal cost of the first unit, c, is not prohibitively high) and Firm 2 absolutely imitates Firm 1 without acquiring any private information of its own. Without the benefit of observing Firm 1's location, Firm 2 would have acquired the same positive quantity of private information that Firm 1 did. But because the observation of Firm 1's location reduces Firm 2's marginal benefit of private information so much that it lies uniformly below marginal cost, Firm 2 never acquires private information after observing Firm 1's location. Firm 2 locates exactly where Firm 1 locates, $y_2^* = y_1^*$, illustrating the case of absolute imitation described in Result 1.

2.3   Aggregate efficiency and absolute imitation

Firm 1 cannot capture the positive informational externality its choice of location provides to Firm 2. To measure the aggregate inefficiency resulting from this informational externality, it is useful to compare aggregate profits between two cases: the decentralized case in which both firms make information and location choices on their own, versus the centralized case in which a central planner simultaneously chooses θ_1, θ_2, y_1 and y_2 to maximize the aggregate profit function $\pi_1(y_1, \theta_1) + \pi_2(y_2, \theta_2)$. In the centralized case where y_1 and y_2 are chosen by the central planner to equal each firm's respective conditional expectation of a, the information acquisition variables θ_1 and θ_2 must be chosen simultaneously to maximize:

$$2\pi_0 - \sigma_a^2(1 - \theta_1) - \sigma_a^2[1 - (\theta_1 + \theta_2 - 2\theta_1\theta_2)/(1 - \theta_1\theta_2)] + c\log(1 - \theta_1) + c\log(1 - \theta_2). \quad (22)$$

The central planner's first-order condition for θ_2 is the same as that faced by Firm 2. Therefore, the central planner also chooses $\theta_2 = 0$ and $y_2 = y_1$. The central planner,

however, chooses a substantially larger quantity of information for Firm 1:

$$\theta_1 = 1 - \frac{1}{2}c/\sigma_a^2 = \frac{1}{2} + \frac{1}{2}\theta_1^*. \tag{23}$$

The expression above relates the central planner's optimal value of θ_1 to Firm 1's own choice in the decentralized regime, showing that the central planner always chooses a larger value. This increase in information raises aggregate profits above the level achieved in the decentralized case by the amount $c(1 + \log(1/2))$. To gauge how large a change in aggregate profits this would be in percentage terms, one needs to refer to aggregate profit in the decentralized case: $2\pi_0 - 2c + c\log(c/\sigma_a^2)$. Thus, the percentage change depends on the magnitude of c relative to π_0, which can be adjusted to achieve arbitrarily large percentage changes, provided $c > 0$, although only as a possibility claim without compelling motivation. These formulas show that the level of change in aggregate profits, as a measure of inefficiency based on the thought experiment of moving from decentralized to centralized regimes, is proportional to the cost-of-information parameter c. Therefore, inefficiency is most severe when information acquisition is expensive, and least severe when it is cheap. This claim depends critically on specification of $C(\theta)$, as the next section shows aggregate efficiency to be nonmonotonic in c when the cost function is quadratic.

2.4 Quadratic cost of information

While the case of absolute imitation nicely represents the real-world phenomenon of imitation and the complete absence of any cost-benefit analysis of unoccupied locations, the intermediate case is interesting as well, where Firm 2 conditions on Firm 1's location but also acquires private information. The remaining analysis in the paper relies on the following quadratic specification of the cost-of-information function:

$$C(\theta) = c\theta^2/2. \tag{24}$$

The key feature of this cost function is that the first unit of information has zero marginal cost, implying that both firms always acquire positive quantities of private information. Solving (13) leads to:

$$\theta_1^* = \sigma_a^2/c, \text{ for } \sigma_a^2 \leq c, \text{ and } 1 \text{ otherwise,} \tag{25}$$

or $\theta_1^* = \min\{\sigma_a^2/c, 1\}$.

With quadratic costs, Firm 2's objective function can be written as:

$$= \pi_0 - (y_2 - \mathrm{E}[a|y_1(\theta_1), x_2; \theta_2])^2 - \sigma_a^2[1 - (\theta_1 + \theta_2 - 2\theta_1\theta_2)/(1 - \theta_1\theta_2)] - c\theta_2^2/2.$$

The first-order condition for θ_2 is:

$$\sigma_a^2(1 - \theta_1)^2/(1 - \theta_1\theta_2)^2 - c\theta_2 = 0. \tag{26}$$

Assuming $\sigma_a^2 \leq c$, one divides (26) through by c and substitutes $\theta_1 = \sigma_a^2/c$, which gives rise to a cubic in θ_2 that turns out to have a unique solution on the unit interval. Following these steps, it is straightforward to re-express (26) using the characteristic equation $h(\theta_2)$:

$$h(\theta_2) \equiv \theta_1^2\theta_2^3 - 2\theta_1\theta_2^2 + \theta_2 - \theta_1(1 - \theta_1)^2 = 0. \tag{27}$$

Because $h(0) = -\theta_1(1 - \theta_1)^2 \leq 0$, and $h(1) = (1 - \theta_1)^3 \geq 0$, there exists at least one solution on the interval $0 \leq \theta_2 \leq 1$. To rule out the possibility of multiple solutions on the closed unit interval, one may examine possible nonmonotonicities of $h(\theta_2)$, which must occur at zeros of the equation:

$$\frac{\partial h(\theta_2)}{\partial \theta_2} = 3\theta_1^2\theta_2^2 - 4\theta_1\theta_2 + 1 = (1 - \theta_1\theta_2)(1 - 3\theta_1\theta_2) = 0. \tag{28}$$

There are two points at which the sign of the curve's slope can change: $\theta_2 = 1/(3\theta_1)$ and $\theta_2 = 1/\theta_1$. The second of these is necessarily to the right of $\theta_2 = 1$, implying that (27) has exactly one solution on the unit interval.

Result 3 (Partial Imitation): Given jointly normal a, x_1 and x_2, and non-decreasing cost function $C(\theta)$ such that $C'(0) < \sigma_a^2(1 - \theta_1^*)$, Firm 2's demand for

16

information is strictly positive, although strictly less than Firm 1's demand for information: $0 < \theta_2^* < \theta_1^*$. In this case, Firm 2's location choice rule can be described as "partial imitation" because Firm 2 chooses a location near, although not exactly the same, as Firm 1. Firm 2's choice of location depends in part on its private signal x_2. Partial imitation implies that y_2 and y_1 are closer than they would be if both firm's relied only on private information.

Figure 1 shows individually profit-maximizing levels of information (i.e., θ_1^* and θ_2^*), chosen by Firms 1 and 2 respectively, for the entire range of (inverse) information costs. The figure also shows aggregate profit, $\pi_1(y_1^*, \theta_1^*) + \pi_2(y_2^*, \theta_2^*)$, in the centralized (topmost curve) and decentralized (second curve from the top) cases. The gap between the two aggregate profit curves, which varies nonmonotonically over the range of information costs, provides one measure of the social cost of imitation. Comparison of the this gap at the extremes versus middle of the range of information costs reveals an interesting nonmonotonicity for the quadratic specification of $C(\theta)$: the social cost of imitation is negligible in environments where information is either very scarce or very abundant, and maximal in the intermediate range of information costs.

This result is different than what was reported above for the exponential cost function. For exponential costs, the cost parameter was bounded from above and, even for cost parameters where neither firm acquired information in the decentralized case, the central planner would always demand a minimum of $\theta_1 = 1/2$. Thus, the gap between centralized and decentralized aggregate profits was maximal where information costs were highest. In contrast, for the case of quadratic costs, if the cost parameter is high enough that neither firm demands information, then it does not pay for the central planner to acquire any information either.

Another interesting feature of Figure 1 is nonmonotonicity of θ_2^*. Whereas Firm 1 always demands more information as c falls, Firm 2's demand for information

can go in either direction in response to a drop in the cost of information. When information is very expensive, neither firm acquires much information and both locate near the unconditional mean μ_a. Because Firm 1's information reduces the marginal benefit of Firm 2's information in all cases, Firm 2 acquires even less information than Firm 1, weighting Firm 1's location more than its own signal. At the other extreme when information is very cheap, Firm 1 acquires so much of the available information about good locations that Firm 2 receives little marginal benefit from its own private information. In this case, the reduction in Firm 2's marginal benefit of information dominates its increase in demand owing to lower acquisition costs.

The variable θ_2^* provides one natural (inverse) measure of the magnitude of imitation, because it represents the extent to which Firm 2 collects private information (i.e., not imitating Firm 1). Alternatively, the magnitude of imitation could be quantified by the squared distance between the firms' locations:

$$(y_2^* - y_1^*)^2 = [\frac{(1 - \theta_1)\theta_2}{1 - \theta_1\theta_2}]^2(x_2 - \theta_1 x_1)^2, \tag{29}$$

which, in expectation, equals:

$$\mathrm{E}[(y_2^* - y_1^*)^2] = \theta_2(1 - \theta_1)^2\sigma_a^2/(1 - \theta_1\theta_2). \tag{30}$$

The distance between the firms' locations is small when θ_2 is near zero or θ_1 is near 1. The expected distance given by the square root of the expression in (30) reaches a maximum of 30 percent of the standard error σ_a. If σ_a is interpreted as the average distance of the ideal location from its unconditional mean, then with this parameterization, firms will on average be closer to each other than the ideal location is from its mean. Another point of interest not directly observable in Figure 1 is that the central planner's solution always prescribes more total information than in the decentralized case. That is, the sum of quantities of information, $\theta_1 + \theta_2$, chosen by the planner is always greater than in the decentralized economy, with a maximum difference of

around 0.35.

For a more detailed view of changes in optimal values of θ_1, θ_2, and aggregate profit moving from the decentralized to centralized regime, Figure 2 shows percentage changes in each of these variables across the same range of inverse information costs. The dotted line at the top of Figure 2 shows percentage change in θ_1, whose uniformly positive sign indicates that the centralized solution always calls for Firm 1 to acquire more information than it chooses on its own. This makes sense, because Firm 1 cannot internalize the benefit it provides to Firm 2 in the decentralized regime. In contrast, Firm 2 usually acquires less information in the centralized regime, but not always. The cases where the central planner dictates that both firms acquire more information correspond to environments in which the cost of information acquisition is relatively large (σ_a^2/c near zero on the x-axis). The range of exogenous parameters in which both firms acquire more private information in the centralized regime reflects synergistic complementarity in the two firms' value of private information, which is nowhere present when using the exponential cost function specification introduced earlier. For quadratic acquisition costs and relatively high values of c, the marginal benefit of Firm 2's private information increases when the central planner raises Firm 1's level of private information.[4] The solid line in Figure 2 is the percentage change in aggregate profit, which is always positive because the central planner optimally utilizes information externalities to achieve greater aggregate profit.

2.5 Discrete choice in acquiring a signal of fixed precision

It is sometimes the case that firms cannot exert continuous control over the precision of private signals they acquire. For example, a consulting firm might offer a report on

---

[4] A related point concerns the nonmonotonicity of Firm 2's response to a change in θ_1. This can be seen analytically in the indeterminate sign of $\frac{\partial \theta_2^*}{\partial \theta_1}$, observed by implicit differentiation of the characteristic equation (27) and noting that $\frac{\partial h(\theta_2)}{\partial \theta_2}$ is positive while $\frac{\partial h(\theta_2)}{\partial \theta_1}$ is of indeterminate sign.

retail sites in a particular city for a fixed price. Similarly, marketing studies priced proportionally to sample size, holding the list of predictors fixed, would not provide purchasers of these services control over R^2. Another example is the decision to spend in-house research time analyzing Census data, which could lead to different model specifications with distinct R^2, but not continuous control over R^2, since the list of Census variables is exogenously fixed.

To investigate the consequences of discretizing the information acquisition decision, this section considers an information market in which both firms make a binary decision of whether to acquire a privately available signal with fixed precision $\bar{\theta}$, at cost $c(\bar{\theta})^2/2$. Note that the error terms in the two firms' private signals are independent, although the signals themselves are of course correlated and the R^2 of each are identical. This implies that, without the signal, Firm 1 faces expected costs of deviating from a equal to σ_a^2. Acquiring the signal, Firm 1 faces expected costs of deviating from a equal to $\sigma_a^2(1 - \bar{\theta})$. Thus, Firm 1's reduction in variance (i.e., increase in expected profit owing to decreased expected deviation from a), achieved as a result of acquiring the signal, equals $\sigma_a^2\bar{\theta}$. Firm 1 acquires the signal if and only if:

$$c(\bar{\theta})^2/2 < \sigma_a^2\bar{\theta}, \quad \text{or} \quad \bar{\theta} < 2\sigma_a^2/c. \tag{31}$$

If Firm 1 acquires the signal but Firm 2 does not, then Firm 2 faces expected costs of deviating from a equal to $\sigma_a^2(1 - \bar{\theta})$. If both firms acquire private signals, then Firm 2's expected cost of deviating from a equals $\sigma_a^2(1 - \bar{\theta})/(1 + \bar{\theta})$. Thus, Firm 2's increase in expected profit by acquiring the signal is $\sigma_a^2\bar{\theta}(1 - \bar{\theta})/(1 + \bar{\theta})$, and it decides to acquire the signal if and only if:

$$c(\bar{\theta})^2/2 < \sigma_a^2\bar{\theta}(1 - \bar{\theta})/(1 + \bar{\theta}), \quad \text{or} \quad \bar{\theta}(1 + \bar{\theta})/(1 - \bar{\theta}) < 2\sigma_a^2/c. \tag{32}$$

Figure 3 shows all possible discrete-information-acquisition environments indexed by two exogenous parameters: the cost of deviating from the ideal location a rel-

ative to the cost of private information acquisition σ_a^2/c, and the fixed precision of information $\bar{\theta}$. In the unshaded region [where both inequalities (31) and (32) fail to hold], neither firm acquires information because the benefit of information relative to its cost is low. In the lightly shaded region [where (31) holds but (32) does not], Firm 1 acquires the private signal and Firm 2 does not, although Firm 2 would have acquired it had Firm 2 been first mover. In the darkly shaded region [where (31) and (32) both hold], both firms acquire private information. Thus, profit maximizing firms in environments with binary choice of acquisition of information may choose to imitate first movers' locations rather than engaging in costly private information acquisition, as was the case in continuous acquisition environments.

3   Discussion

Rather than evaluating shortcuts, or heuristics, according to context-free domain-general normative criteria, the aim of this paper is to explain when the imitation heuristic works well in terms of social efficiency and where it is, unfortunately, mismatched to the decision environment, contributing to persistently under-developed regions in central cities. There is a long and distinguished literature on spatial agglomerations of people and commerce (Christels, 1933; Lösch, 1938; Zipf, 1949; Berry, 1961). Economists have advanced formal models of spatial organization, from Hotelling (1929) to Krugman (1993), and beyond.[5] Economists have also contributed a rich theoretical literature explaining why imitation is individually advantageous in various settings (Sinclair, 1990; Welch, 1992; Banerjee, 1992; Bikhchandhani, Hirshleifer and Welch, 1992; Ellison and Fudenberg, 1995; Vega-Redondo, 1997;

---

[5]Boschma and Frenken (2007) provide a lucid discussion of economic geographys institutional focus in contrast with new economic geographys neoclassical methodology, which attempts to explain uneven distributions of economic activity in terms of universal processes driven by mobile factors of production.

Schlag 1998, 1999; Offerman, Potters, and Sonnemans, 2002; Apesteguia, Huck, and Oechssler, 2003; Dutta and Prasad, 2004; Anderson, Ellison and Fudenberg, 2005), with interesting empirical applications as well.[6] There is, however, very little in the economics literature marrying these two themes of imitation and location choice.[7]

Venerable lines of economic research develop models of location choice (Hotelling, 1929; Alonso, 1964; Muth, 1968) as well as spatial agglomerations and their statistical determinants (Konishi, 2005; Kobrin, 1985). The issue of ethnic enclaves and spatial patterns resulting from individual decisions about where to move also relates to the problem of imitation in location choice, at least from an abstract modeling perspective (Gross and Schmitt, 2000; Huff, 1962). Not all those who analyze spatial patterns focus on processes of agglomeration. Some argue that Hotelling-type economies should produce dispersion rather than concentration (d'Aspremont and Gabszewicz, 1979). Similarly, Kain (1968) focuses on decentralization (i.e., the undoing of spatial agglomerations) and the unequal impacts of suburbanization on labor market opportunities for blacks and other ethnic minorities. According to Glaeser, Hanushek and Quigley (2004), Kain's spatial explanations for persistently high unemployment in minority neighborhoods played a large role in directing the attention of economists to disparities based on race and ethnicity.

---

[6]Rodgers (1952) describes dramatic spatial concentrations of steel production in the U.S., and the possibility that these concentrations might undermine national security. Similarly, Rees (1978) describes spatial concentrations in the rubber industry. Mansfield (1961) provides empirical evidence linking firms' decisions to introduce new techniques of production to the proportion of firms already using that technique, in line with widely used gravity models in the social sciences. Geertz (1978) observes spatial agglomeration according to product type in bazaars in Algeria. Walcott (1990) finds agglomerations of biotech firms in Atlanta suggestive of imitation as a strategy for coping with scarcity of information. Fairen (1996) argues that imitative behavior may best explain why automobile manufacturers produce very similar models of cars. And Seamans (2006) investigates spatial clustering in the cable television industry.

[7]One exception is the experimental Hotelling economy analyzed by Camacho-Cuena et al (2005), demonstrating spatial agglomerations in the lab, but not always as the result of decision-making processes that follow the standard model. The international finance literature, too, frequently studies interdependencies among firms' investment decisions (Kindleberger, 1983), and imitation is an established theme in international trade (Schmitt, 1995).

In contrast to economics, the administrative management literature has devoted considerable attention to imitation in location choice (Guillen, 2002; Haunschild, 1993; Haveman, 1993). Descriptive models in this literature focus on how exit and entry of other firms allow managers to make inferences about expected levels of profitability, leading to correlated entry and exit decisions across firms (Baum, Li and Usher, 2000; Miner and Haunschild, 1997), which is consistent with the formal model presented in this paper. Another reason for imitation mentioned in this literature is sociological imperatives to adhere to norms (Abrahamson and Rosenkopf, 1993), according to which it pays to imitate peer decisions even in the absence of internal reasons for adopting strategies that peer firms have adopted. Legitimacy is put forward as still another reason why managers may eschew independent approaches in favor of imitation of peers whose actions are perceived as legitimate (DiMaggio and Powell, 1983; Haunschild and Miner, 1997; Fligstein, 1985). An interesting aspect of this literature is that, in many of these models, the more predictable the environment, the stronger the incentive to imitate (Argote, Beckman and Epple, 1990), which is opposite of the model in this paper. There is also a vast operations research literature applying constrained optimization to spatial decision problems.

Empirical accounts from interview studies (Schwartz, 1987, 2004; Bewley, 1999; Berg, 2007) favor the position that firms rely on simplifying rules of thumb, or heuristics. Wiessbourd (1999) reports that businesses in Chicago use simple rules of thumb to decide on locations, which work well in environments with lots of information, but tend to reinforce negative stereotypes and leave profitable opportunities unexploited in low information environments. Anecdotal evidence corroborates the large role of imitation put forward in this paper. For example, according to one individual involved in location decisions for the German discount supermarket chain Lidl, its location decisions follow a simple rule of thumb: build a store wherever Aldi, Lidl's

primary competitor, has a location (Scheibene, 2007, personal communication).

Given that firms in the real world condition location decisions on observed locations of others, it is natural to investigate whether imitation can be rationalized within the profit maximizing framework. The model presented in Section 2 tackles this problem and the more important issue, from a policy perspective, of the aggregate consequences of shortcut decision rules, or heuristics. The central question is whether under-utilization of urban resources results from a process in which firms consider employing those resources and wind up deciding they are unprofitable, or whether the imitation heuristic, perhaps suitable for some investment environments, leads to inefficient clustering in favored suburban locations and systematic neglect of entire regions within cities.

The magnitude of the abandoned property problem in US central cities is apparent in the economics literature as early as the 1970s (Stegman and Rasmussen, 1980). Whereas the model of Caplin and Leahy (1998) provides a general rationalization for abandonment or under-utilization, the model in this paper has the advantage of identifying conditions in the external environment necessary for tension to exist between the individual advantages and social consequences (in terms of aggregate economic efficiency) of imitation.

4 Conclusion

The model presented in this paper draws on empirical accounts of spatial concentration in well-established retail centers of affluent suburbs by firms that inefficiently overlook profitable opportunities in urban neighborhoods (Berg, 2007; Helling and Sawicki, 2003; Sabety and Carlson, 2003; Francica, 2000; Weissbourd, 1999). Seeking a better understanding of opportunistic information sharing that leads to imitation,

the model demonstrates that imitation is consistent with expected profit maximization, although it usually results in lower aggregate profit than would be achieved by a benevolent central planner. The benchmark relative to which efficiency is measured does not assume a new institution for pooling information resources. Rather, it takes the sequence of moves as given and simultaneously chooses all firms' choice variables (i.e., Firm 1 and Firm 2's quantities of private information acquisition and choices of location) to maximize aggregate profit. Firms are assumed to be identical except that later movers can make use of earlier movers' locations as an additional conditioning variable.

As emphasized already, imitation in location choice is not uniformly bad for aggregate efficiency. There is a genuine positive externality flowing from early to later movers. Imitation usually helps exploit this positive externality to some extent, but not as fully as aggregate efficiency requires. As shown in Figure 2, over most of the range of the cost-of-information parameter c, the central planner prescribes more information for Firm 1 and less for Firm 2. Thus, in this parameter range, the central planner fully exploits the positive externality by increasing, not reducing, the extent of imitation. When c is very large, however, the central planner requires that both firms acquire additional private information, implying that a reduction in imitation is needed to achieve social efficiency. This case might argue for public provision of neighborhood-level demographic and crime information that can be used to estimate revenues and costs, or perhaps direct subsidies for first movers into neighborhoods seeking (re-)development.

Thus, whether imitation is consistent with the greater good depends on the informational environment. It is interesting to consider what the model predicts for well-established suburban areas with thriving concentrations of retail compared with low-income neighborhoods without much commercial activity at all. Insofar as sub-

urbs enjoy well-defined land use rules and relatively liquid markets for efficiently channeling development capital to profitable locations, the suburban environment is informationally easy, in the sense that it is cheap to discover where profits are earned. The model says that, in an informationally easy environment like that, the imitation heuristic is both individually effective and socially useful. In contrast, when information is expensive or scarce, imitation remains individually effective yet undermines social efficiency.

A general feature of the model is that firms would always prefer to be second mover and, if the cost of time is low enough, would choose to wait rather than move first, consistent with the idea of spatial lock-in and systematic neglect of ignored resources in central cities. Second movers always enjoy greater expected profit because freely available observation of the first mover's location results in lower total information costs over the entire parameter space. This suggests a motive for first movers to try to be so large that they exhaust all monopoly rents associated with a particular location, perhaps applicable to the phenomenon of big box retail and highly coordinated development that involves a few very large first movers rather than a long sequence of smaller movers.

Policies aimed at spurring business development in poor neighborhoods typically rely on the standard economic model of profit-maximization and its assumption that firms consider vast, if not infinite, sets of alternatives before choosing where to invest. Empirical work, however, points to strong limitations on firms' consideration sets and their ability to make reliable spatial predictions in terms of profit. Very different policy approaches are called for if firms' decision processes diverge from the standard model and are better represented by a simple decision tree that eliminates neighborhoods from consideration based on a single reason—for example, because there are no other firms there, or because of statistically unsubstantiated fears about

26

high crime, or because of managers' inherent preferences for areas that are personally familiar to them. Future work detailing the size and sources of firms' consideration sets when making location decisions, and improved description of the actual decision process used to choose an element from the consideration set, would be useful.

Milton Friedman's as-if methodology argues that it is acceptable to use an incorrect model of consumers and firms' decision processes as long as it predicts behavior accurately. As-if models may be misguided, though, for purposes of designing policies to modify behavior. As every student of statistics learns, one can always add parameters to achieve arbitrarily good levels of fit with respect to a particular data set. But fit is not prediction. Prediction requires generalization out-of-sample, often beyond the current parameter range. Insofar as policy makers working on urban development wish to achieve outcomes significantly outside the status-quo range of behaviors, they will require empirically grounded behavioral models with sufficient psychological realism to achieve accurate descriptions of firms and consumers' decision processes.

References

Abrahamson, E., and Rosenkopf, L. (1993). Institutional and competitive bandwagons: Using mathematical modeling as a tool to explore innovation diffusion. Academy of Management Review, 18, 487-517.

Alonso, W. (1964). Location and land use: Toward a general theory of land rent. Cambridge, MA: Harvard University Press.

Anas, A., and Rhee, H. J. (2007). When are urban growth boundaries not second-best policies to congestion tolls?. Journal of Urban Economics 61(2), 263-286.

Anderson, R. M., Ellison, G., and Fudenberg, D. (2005). Location choice in two-sided markets with indivisible agents. HIER Working Papers 2056, Harvard University.

Andersson, F., Burgess, S., and Lane, J. I. (2007). Cities, matching and the productivity gains of agglomeration. Journal of Urban Economics, 61(1), 112-128.

Apesteguia, J., Huck, S., and Oechssler, J. (2003). Imitation—Theory and experimental evidence. CESifo Working Paper Series No. 1049, University of Munich.

Argote, L., Beckman, S. L., and Epple, D. (1990). The persistence and transfer of learning in industrial settings. Management Science, 36, 140-154.

Baum, J., Li, S. X., and Usher, J. (2000). Making the next move: How experiential and vicarious learning shape the locations of chain's acquisitions. Administrative Science Quarterly, 45(4), 766-801.

Berg, N. (2007). Business development decisions in Dallas: Assessing the need for new tools in new environments. Working Paper, University of Texas-Dallas.

Berry, B. J. L. (1961). City size distributions and economic development. Economic Development and Cultural Change, 9, 573-588.

Bewley, T. F. (1999). Why wages don't fall during a recession, Cambridge, MA: Harvard University Press.

Bikhchandani, S., Hirshleifer, D., and Welch, I. (1998) Learning from the behavior of others: Conformity, fads and information cascades. Journal of Economic Perspectives, 12(3), 151-170.

Bosch-Domnech, A., and Vriend, N. J. (2003). Imitation of successful behavior in Cournot markets. The Economic Journal, 113, 495-524.

Boschma, R. A., and Frenken, K. (2007). Applications of evolutionary economic geography, In Frenken, K. (Ed.), Applied Evolutionary Economics and Economic Geography. Cheltenham: Edward Elgar.

Boulding, K. (1996). Economic Analysis. Vol. I, Microeconomics. 4th Ed. New York: Harpers.

Bray, T. (2007). Dallas is not as dangerous as you think: The crime we have is bad, but it's concentrated, Dallas Morning News, April 6, Viewpoints.

Brueckner, J. K., and Rosenthal, S. S. (2005). Gentrification and neighborhood housing cycles: Will America's future downtowns be rich?. CESifo Working Paper Series No. 1579. Accessed October 28, 2007, at http://ssrn.com/abstract=852525.

Camacho-Cuena, E., Garca-Gallego, A., Georgantzs, N., and Sabater-Grande, G. (2005). Buyerseller interaction in experimental spatial markets. Regional Science and Urban Economics, 35, 89-108.

Chang, S. J. (1995). International expansion strategy of Japanese firms: Capability building through sequential entry. Academy of Management Journal, 38, 383-407.

Chang, S. J., and Rosenzweig, P. M. (2001) The choice of entry mode in sequential foreign direct investment. Strategic Management Journal, 22, 747-776.

Christaller, W. (1933). Die zentralen Orte in Sddeutschland. Jena, Germany. Gustav Fischer. [trans. Baskin, C. W. (1996). Central Places in Southern Germany. Englewood Cliffs, NJ: Prentice Hall].

Chung, W. (2001). Mode, size, and location of foreign direct investments and industry markups. Journal of Economic Behavior and Organization, 45, 185-211.

Cyert, R. M., and March, J. G. (1963). A behavioral theory of the firm, 2nd Ed. Cambridge, MA: Blackwell Business.

d'Aspremont, C., Gabszewicz, J. J. and Thisse, J. F. (1979). On Hotelling's stability in competition. Econometrica, 47, 1145-1150.

DiMaggio, P. J., and Powell, W. W. (1983). The iron cage revisited: Institutional isomorphism and collective rationality in organizational fields. American Sociological Review, 103, 1-37.

Dudey, M. (1990). Competition by choice: The effect of consumer search on firm location decisions. American Economic Review 80, 10921104.

Dudey, M. (1993). A note on consumer search, firm location choice, and welfare. Journal of Industrial Economics, 41, 323-331.

Dutta, J., and Prasad, K. (2004). Imitation and long run outcomes. Berkeley Electronic Press, 4(1), 1163-1163.

Eaton, B. C., and Lipsey, R. G. (1976). The non-uniqueness of equilibrium in the Löschian location model. American Economic Review, 66, 77-93.

Eaton, B. C., and Lipsey, R. G. (1979). Comparison shopping and the clustering of homogeneous firms. Journal of Regional Science, 19, 421435.

Eeckhoudt, L., Godfroid, P., and Gollier, C. (2001). Multiple risks and the value of information, Economic Letters, 73, 359-365.

Ellison, G., and Fudenberg, D. (1995). Word of mouth communication and social learning. Quarterly Journal of Economics, 110, 93-126.

Fairen, V. (1996). A replicator theory model of competition through imitation in the automobile market. Journal of Economic Behavior and Organization, 29, 141-157.

Fischer, J. H., and Harrington, J. E. Jr. (1996). Product variety and firm agglomeration. The RAND Journal of Economics, 27, 281-309.

Fligstein, N. (1985). The spread of the multidivisional form among large firms, 1919-1979. American Sociological Review, 50, 377-391.

Francica, J. (2000). Location analysis tools help Starbucks brew up new ideas. Business Geographics, 8, 3233.

Geertz, C. (1978). The bazaar economy: Information and search in peasant marketing. American Economic Review, 68, 28-32.

Gigerenzer, G., and Selten, R. (Eds.) (2002). Bounded rationality: The adaptive toolbox. Cambridge, MA: MIT Press.

Gigerenzer, G., Todd, P. M., and The ABC Research Group. (1999). Simple heuristics that make us smart. New York: Oxford University Press.

Glaeser, E., Hanushek, E., and Quigley, J. (2004). Opportunities, race, and urban location: The influence of John Kain. HIER Working Paper 2030, Harvard University.

Gross, D., and Schmitt, N. (2000). Do birds of a feather flock together? Immigration flows and cultural clustering in host countries. Working Paper, Simon Fraser University.

Guillen, M. (2002). Structural inertia, imitation and foreign expansion: South Korean firms and business groups in China, 1987-1995. Academy of Management Journal 2(4), 389-421.

Haunschild, P. R. (1993). Interorganizational imitation: The impact of interlocks on corporate acquisition activity. Administrative Science Quarterly, 38, 564-592.

Haveman, H. A. (1993). Follow the leader: Mimetic isomorphism and entry into new markets. Administrative Science Quarterly, 38, 564-592.

Helling, A., and Sawicki, D. (2003). Race and residential accessibility to shopping

and services. Housing Policy Debate, 14, 69101.

Helsley, R. W., and Strange, W. C. (1999). Gated communities and the economic geography of crime. Journal of Urban Economics, 46(1), 80-105.

Helsley, R. W., and Strange, W. C. (2005). Mixed markets and crime. Journal of Urban Economics, 89(7), 1251-1275.

Hilton, R. (1981). Determinants of information value. Management Science, 27, 5764.

Hotelling, H. (1929). Stability in competition. Economic Journal, 39, 41-57.

Huff, D. L. (1962). A note on the limitations of interurban gravity models. Land Economics, 38, 64-66.

Hutchinson, J. M. C. (2005). Is more choice always desirable? Evidence and arguments from leks, food selection, and environmental enrichment. Biological Reviews, 80, 7392.

Kain, J. F. (1968). Housing segregation, Negro employment and metropolitan decentralization. Quarterly Journal of Economics, 82, 175-197.

Kanemoto, Y. (1990). Optimal cities with indivisibility in production and interactions between firms. Journal of Urban Economics, 27, 46-59.

Kindleberger, C. P. (1983). International banks as leaders or followers of international business. Journal of Banking and Finance, 7, 583-595.

Kobrin, S. J. (1985). Diffusion as an explanation of oil nationalization (or the domino effect rides again). Journal of Conflict Resolution, 29, 3-32.

Kogut, B. (1983). Foreign direct investment as a sequential process. In Audretch, D. and Kindleberger, C.P. (Eds.), The Multinational Corporation in the 1980s (pp. 38-56). Cambridge, MA: M.I.T. Press.

Konishi, H. (2005). Concentration of competing retail stores. Journal of Urban Economics, 58, 488-512.

Krugman, P. (1993). On the number and location of cities. European Economic

Review. 37, 293-298.

Lösch, A. (1938). The nature of economic regions. Southern Economic Journal, 5, 71-78.

Mansfield, E. (1961). Technical change and the rate of imitation. Econometrica, 29, 741-766.

March, J. G. (1988). Decisions in organizations. New York: Blackwell.

Miner, A. S., and Hanuschild, P. R. (1997). Modes of interorganizational imitation: The effects of outcome salience and uncertainty. Administrative Science Quarterly, 42, 472 500.

Muñiz, I., and Galindo, A. (2005). Urban form and the ecological footprint of commuting: The case of Barcelona. Ecological Economics, 55, 499-514.

Muth, R. (1968). Cities and housing. Chicago: University of Chicago Press.

Noble, J., Todd, P. M., and Tuci, E. (2001). Explaining social learning of food preferences without aversions: An evolutionary simulation model of Norway rats. Proceedings of the Royal Society of London B: Biological Sciences, 268 (1463), 141-149.

Offerman, T., Potters, J., and Sonnemans, J. (2002). Imitation and belief learning in an oligopoly experiment. Review of Economic Studies, 69, 973-997.

Prescott, E. C., and Visscher, M. (1977). Sequential location among firms with foresight. The Bell Journal of Economics, 8, 378-393.

Rees, J. (1978). On the spatial spread and oligopolistic behavior of large rubber companies. Geoforum, 9, 319-30.

Rodgers, A. (1952). Industrial inertia—A major factor in the location of the steel industry in the United States. Geographical Review, 42, 56-66.

Rosenthal, S. S., and Strange, W. C. (2001). The determinants of agglomeration. Journal of Urban Economics, 50(2), 191-229.

Sabety, P., and Carlson, V. (2003). Using information to drive change: New ways of moving markets. Washington, DC: Brookings Institution.

Schlag, K. (1998). Why imitate, and if so, how? A boundedly rational approach to multi-armed bandits. Journal of Economic Theory, 78, 130-56.

Schlag, K. (1999). Which one should I imitate? Journal of Mathematical Economics, 31, 493-522.

Schmitt, N. (1995). Product imitation, product differentiation and international trade. International Economic Review, 36, 503-608.

Schwartz, H. (1987). Perception, judgment and motivation in manufacturing enterprises: Findings and preliminary hypotheses from in-depth interviews. Journal of Economic Behavior and Organization 8, 543-65.

Schwartz, H. (2004). The economic analysis underlying corporate decision making: What economists do when confronted with business realities. Business Economics, 39, 50-59.

Seamans, R. (2006). How does clustering by cable TV firms affect technology diffusion? Working Paper, University of California-Berkeley.

Simon, H. A. (1954). Bandwagon and underdog effects and the possibility of election predictions. Public Opinion Quarterly, 18, 245-253.

Simon, H. A. (1955). A behavioral model of rational choice. The Quarterly Journal of Economics, 69, 99-118.

Sinclair, P. J. N. (1990). The economics of imitation. Scottish Journal of Political Economy, 37, 113-144.

Turner, M.A. (2007). A simple theory of smart growth and sprawl. Journal of Urban Economics, 61, 2144.

Vega-Redondo, F. (1997). The evolution of Walrasian behavior. Econometrica, 65, 375-384.

Verdier, T., and Zenou, Y. (2004). Racial beliefs, location, and the causes of crime. International Economic Review, 45(3), 731-760.

Walcott, S. M. (1999). High tech in the deep south: Biomedical firm clusters in metropolitan Atlanta. Growth and Change, 30(1), 48-74.

Weissbourd, R. (1999). The market potential of inner-city neighborhoods: Filling the information gap. Washington, DC: The Brookings Institution.

Willinger, M. (1989). Risk aversion and the value of information. Journal of Risk and Insurance, 56(2), 32028.

Wolinsky, A. (1983). Retail trade concentration due to consumers' imperfect information. Bell Journal of Economics, 14, 275-282.

Zipf, G. K. (1949). Human behavior and the principle of least effort. Cambridge, MA: Harvard University Press.

Figure 1: Centralized aggregate profit envelope (dash-dotted line), aggregate profit under individual choice (solid line), Firm 1's (dotted) and Firm 2's (starred) private information acquisition, as a function of inverse information cost (from most to least expensive)
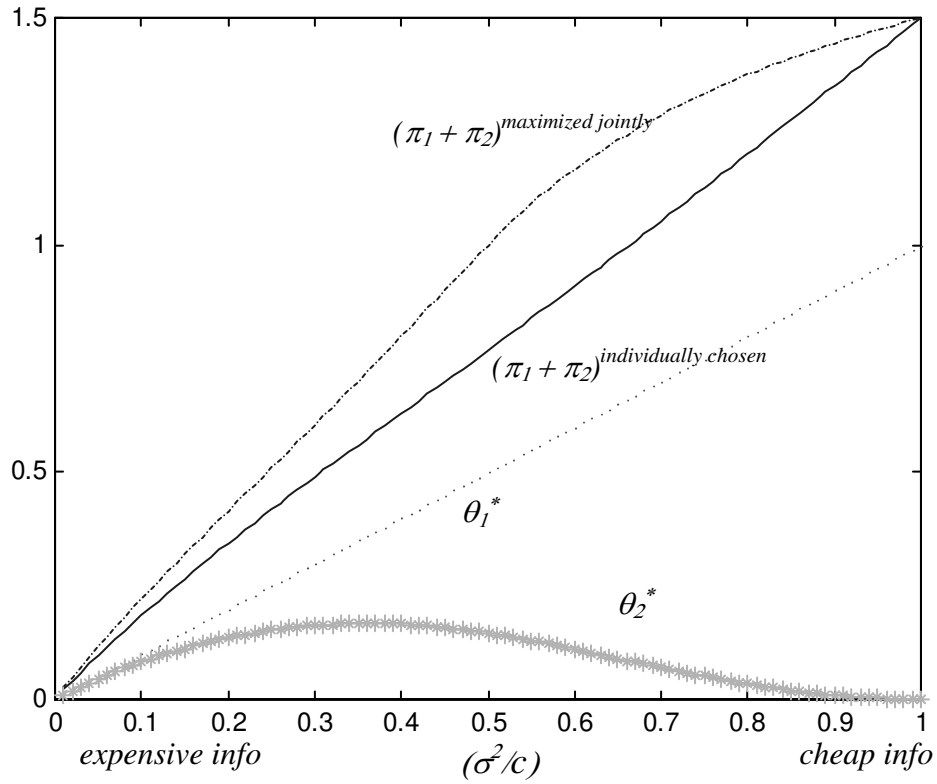

$(\pi_1 + \pi_2)^{maximized\ jointly}$

$(\pi_1 + \pi_2)^{individually\ chosen}$

θ_1^*

θ_2^*

expensive info

(σ^2/c)

cheap info

Figure 2: From decentralized to centralized regimes: Percentage change in aggregate expected profit (solid line), Firm 1's information acquisition (dotted), and Firm 2's information acquisition (starred), as a function of inverse information cost
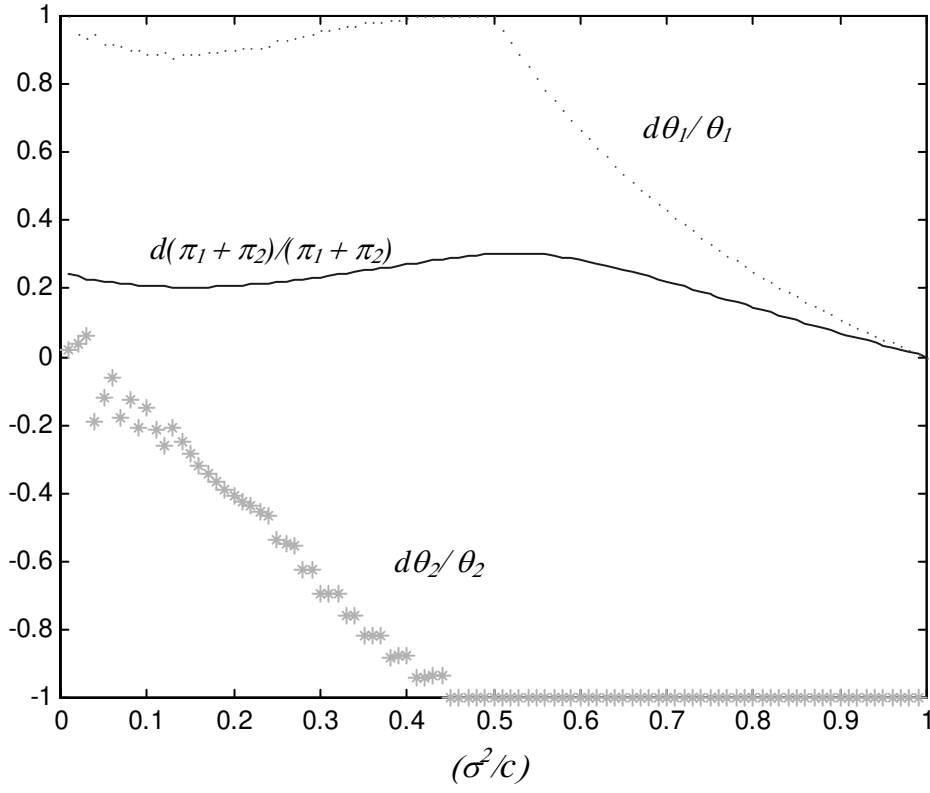
Figure 3: Binary information acquisition decisions as a function of inverse cost of information (x-axis) and the fixed precision of the private signal (y-axis): Environments in which neither firm acquires private information (unshaded), Firm 1 acquires private information and Firm 2 does not (lightly shaded), and both firms acquire private information (darkly shaded)
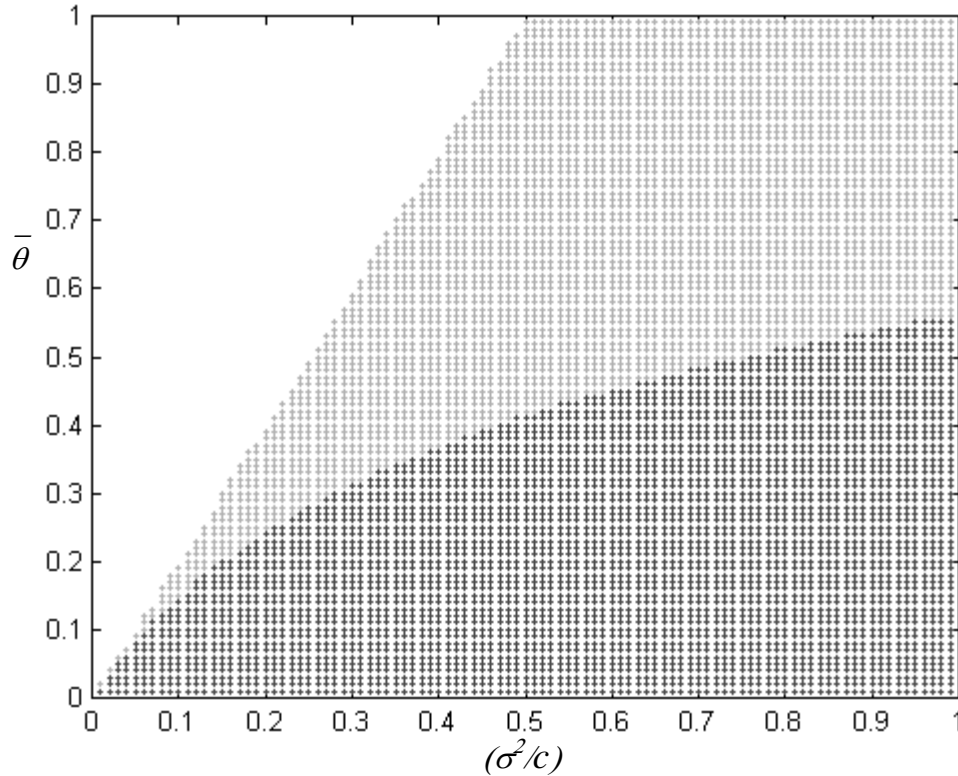

Figure 3 shows a partition of the universe of possible environments, indexed by the inverse relative cost of information and fixed precision of the available signal. Firms make discrete choices of whether to acquire the signal. Whether firms acquire the signal depends on the signal's cost, its (fixed) quality, and--most importantly--whether the firm is first mover or not. In the unshaded region, neither firm acquires information because costs of information are high relative to the cost savings in deviating from the ideal location a. In the darkly shaded region, both firms acquire information because information costs are relatively low. In the lightly shaded region, however, Firm 1 acquires information but Firm 2 does not. This difference is purely the result of first- and second-mover status, as Firm 2 would have acquired the signal had it not been able to freely extract information by observing Firm 1's location.