RAND Journal of Economics
Vol. 55, No. 2, Summer 2024
pp. 169–198

The timing and location of entry in growing markets: subgame perfection at work

Bryan Bollinger*

Ulrich Doraszelski**

Kenneth L. Judd***

and

Ryan C. McDevitt‡

We develop and analyze a dynamic model in which firms decide when and where to enter a growing market. We do not pre-specify the order of entry, allowing instead for the leader and follower to be determined endogenously. We characterize the subgame perfect equilibria of the dynamic game and show the times and locations of entry are governed by the threat of preemption, which leads to premature entry, less extreme locations, and the dissipation of rents. Using data on gas stations, restaurants, and hotels in isolated markets, we find results consistent with subgame perfection for gas stations and three-star hotels.

1. Introduction

■ Consider a growing market. At each point in time, a firm must decide whether to enter and, if so, where to position itself within the market. Any gain from being the first entrant generates incentives for firms to preempt each other. In this article, we account for the threat of preemption by developing a dynamic model of entry around a model of spatial competition. We characterize the subgame perfect equilibria (SPE) and show that the threat of preemption leads to premature entry, less extreme locations, and the dissipation of rents. In short, preemption matters.

Building on Hotelling (1929), the literature on spatial competition attempts to explain an industry's structure by analyzing a static game of location choice followed by price competition. A static game implicitly assumes an unchanging environment, in particular by holding market

* Stern School of Business, New York University.
** Wharton School, University of Pennsylvania.
*** Hoover Institution, Stanford University.
‡ Fuqua School of Business, Duke University; ryan.mcdevitt@duke.edu.
Thanks to Michaela Draganska, Dino Gerardi, and Christian Yankov for their excellent feedback. Many thanks also to Nah Lee for excellent research assistance.
ryan.mcdevitt@duke.edu

size fixed. As a result, there is a fixed number of firms in the market. Yet, an industry's structure changes with market size. Although the market is small, it supports at most one firm. As the market grows over time, another firm may enter. In other words, the industry starts with a monopoly phase which may be followed by a duopoly phase. Fully understanding when firms choose to enter the market and where they position themselves within it requires a dynamic model.

Adding a time dimension to the model can overturn the principle of maximum differentiation that is central to the literature on spatial competition (see d'Aspremont et al. (1979), Economides (1984, 1986), and Neven (1985), among others). In a static game of location choice followed by price competition, a firm has an incentive to differentiate itself from its rival to soften price competition. In the extreme, this implies that firms locate at opposite ends of the market. Adding a time dimension to the model can overturn the principle of maximum differentiation because a more central location benefits the first entrant in two ways. First, a more central location increases the incumbent's profits during the monopoly phase. Second, a more central location decreases the potential entrant's profits during the duopoly phase, thereby delaying further entry.[1] The first entrant may therefore choose to position itself at or near the center of the market rather than at its extremes.

Although a small literature considers the timing decision alongside the location decision (Zhou and Vertinsky, 2001; Lambertini, 2002), it has arguably failed to capture the essence of a fully dynamic entry process because the employed equilibrium concept, by pre-specifying the order of entry, rules out the threat of preemption. At the outset of the game, firms commit, one by one, to a time and location of entry.[2] As pointed out by Fudenberg and Tirole (1985) in their seminal critique of the models of technology adoption in Reinganum (1981b, 1981a), such a sequential pre-commitment equilibrium (SEQPE) fails to be subgame perfect. To see this, consider a SEQPE in which firm 1 plans to enter the market at time t_1 and location x_1 and firm 2 at some later time $t_2 > t_1$ and location x_2. We show below that the leader's payoff exceeds the follower's payoff. Firm 2 therefore has an incentive to preempt firm 1 by entering sightly before t_1 at location x_1, thereby securing itself a payoff arbitrarily close to that of firm 1. Because the follower has an incentive to become the leader, the SEQPE is not subgame perfect, and the threat of preemption is operative until payoffs equalize.[3]

In short, if firms pre-commit to a time and location of entry, preemption is ruled out by assumption. Although this may be appropriate in some applications, in many others, we expect preemption to be critically important for industry structure and dynamics. To account for the threat of preemption, we specify a fully dynamic entry process and characterize the SPE of the dynamic game. In contrast to a SEQPE, the roles of leader and follower are determined endogenously in a SPE. We show that the threat of preemption governs both the times and the locations of entry in a growing market.

By considering the timing decision together with the location decision, we move beyond recent theoretical work that highlights preemption and its implications for the timing decision, such as Smirnov and Wait (2015), who develop a dynamic model with two firms and characterize the SPE, and Shen and Villas-Boas (2010), who develop a dynamic model of a growing market and show that the ability of early entry to deter future competitors' entry leads firms to enter the market at a rate faster than demand is expanding.

[1] The observation that a more central location can deter entry has been made before by Neven (1987) in a version of the basic game where location choices are made sequentially rather than simultaneously. Because there is no time dimension in the model, there are no separate monopoly and duopoly phases.

[2] Lambertini (2002) assumes that the follower behaves non-strategically and enters at a fixed time. He shows that the later the follower enters, the closer to the center of the market the leader positions itself. Because there is no fixed cost of entry, the leader always enters immediately.

[3] Riordan (1992) incorporates asymmetries into the model of Fudenberg and Tirole (1985) and shows that, with two firms, the more efficient firm enters first. Argenziano and Schmidt-Dengler (2013) extend Fudenberg and Tirole (1985) to more than two firms and show that the time of first entry in a duopoly is a lower bound on the time of first entry in any oligopoly and that more firms may delay first entry.

In addition to incorporating the time and location dimensions of the entry decision, we establish testable predictions that allow us to distinguish between SEQPE and SPE in the data. Specifically, we show that, conditional on the location of first entry, the rate of market growth does not affect the threshold for market size at the time of first entry under SEQPE, whereas the threshold is a decreasing function of the rate of market growth under SPE. In our empirical application, we study the timing and location decisions of the first entrant for gas stations, restaurants, hotels, and three-star hotels in geographically isolated markets around highway exits and intersections, a nearly ideal setting for testing whether SEQPE or SPE better fits the data: A firm's location has a direct effect on its profits as traffic patterns favor firms in closer proximity to highway exits, whereas firms of the same type that are close to each other may cannibalize sales and incite price competition. Furthermore, highway traffic data provide good measures of market size and the rate of market growth.

We find that gas stations are consistent with SPE but restaurants and hotels are not. This finding may be explained by institutional details, in particular the fact that restaurants and hotels combine spatial differentiation with other dimensions of product differentiation (Mazzeo, 2002; Butters and Hubbard, 2023). Indeed, when we isolate spatial differentiation by examining three-star hotels, the data are again consistent with SPE. Taken together, we conclude that SPE can be a more natural equilibrium concept than SEQPE for understanding the evolution of growing markets.

We view our approach of testing an implication of subgame perfection as a complement to the structural approaches used to detect preemption and measure its implications. In contrast to the theoretical literature, the empirical literature on preemption is small, with notable contributions by Schmidt-Dengler (2006), Igami and Yang (2016), Zheng (2016), and Fang and Yang (2019, 2022). Schmidt-Dengler (2006) studies the adoption of MRI scanners by hospitals. Following Fudenberg and Tirole (1985) and the ensuing theoretical literature, he defines preemption by contrasting SEQPE and SPE in a similar way as we do. One contribution of our article vis-a-vis Schmidt-Dengler (2006) is that we consider the timing decision together with the location decision and show that the threat of preemption spills over from the former into the latter.

Later articles depart from Fudenberg and Tirole (1985) to define preemption. Igami and Yang (2016) consider the decisions of hamburger chains regarding the number of outlets in a market. Focusing on Markov perfect equilibria (MPE), and thus a subset of SPE, they say that McDonald's has a motive to preempt to the extent that its entry probability in a market is lower in a counterfactual than in the MPE, where the counterfactual aims to force McDonald's rivals to behave as if the number of McDonald's outlets in the market does not matter. Also restricting attention to the timing decision, Fang and Yang (2019, 2022) identify preemption by decomposing a firm's marginal benefit of entry in the equilibrium conditions, where shutting down the relevant term in a counterfactual allows them to measure the implications of preemption. Fang and Yang (2019) use this approach to study the entry decisions of fast casual taco chains, whereas Fang and Yang (2022) study the decisions of coffee chains regarding the number of outlets in a market. In addition, Zheng (2016) considers the entry decisions of big box retail chains and defines preemption as a one-period deviation from equilibrium; because a retail chain makes decisions simultaneously for all markets, it decides on both the timing and location of entry, similar to our article.

Our approach of testing an implication of subgame perfection does not allow us to conduct counterfactuals and thus limits what we can learn about the industry being studied. On the other hand, it incorporates the time and location dimensions of the entry decision and may be less dependent on the details of the model and the simplifying assumptions added to make the model computationally tractable.[4]

[4] For example, Zheng (2016) truncates the time horizon at the end of her sample period and assumes that firms move in alternating periods using a two-stage budgeting process in which they first allocate financial or managerial resources to groups of markets and then, conditional on this budget allocation, make entry decisions.

Our article is more generally related to recent empirical applications of the Ericson and Pakes (1995) framework, including Arcidiacono et al. (2016), Collard-Wexler (2013), and Sweeting (2013). Although the threat of preemption is of course operative in any MPE, these articles make no attempt to isolate it. Arcidiacono et al. (2016) and Collard-Wexler (2013) consider the entry and exit decisions of retail chains and ready-mix concrete plants, respectively, assuming that geographical markets are independent of one another.[5] Given this assumption, firms (and the analyst) can treat each market separately and only the time dimension of the entry decision matters. Whereas Sweeting (2013) maintains that geographical markets are independent of one another, in his model, each firm decides on the programming format for each of the radio stations it owns. Hence, in a given period, a radio station can be out of the market, or if it is in the market, in different locations in product space, and a firm's decisions regarding the time and location of entry of the radio stations it owns are influenced by the threat of preemption.

The remainder of the article is organized as follows. Section 2 introduces the model. Section 3 characterizes the SEQPE, and Section 4 the SPE. Section 5 sets up a welfare benchmark by considering an omnipotent social planner that controls the time and location of entry as well as firms' pricing decisions in the product market. Section 6 uses numerical analysis to compare the outcome of the game under SEQPE and SPE and assess their welfare implications. We show that the threat of preemption leads to premature entry, less extreme locations, and the dissipation of rents. The degree of rent dissipation may be far greater under SPE than under SEQPE, and we provide an example where up to 237 times as much surplus is wasted under SPE than under SEQPE. Section 7 contains our empirical application. Section 8 concludes. Online Appendix A contains details on price competition; online Appendices Appendix B, C, and D most proofs; and online Appendix E details on the data we use in our empirical application.

2. Model

■ There are two firms $i \in \{1, 2\}$. Firm i decides a time $t_i \geq 0$ and a location $x_i \in [0, 1]$ to enter the market. Firms incur a fixed cost of entry $F > 0$. Time is continuous and the horizon is infinite. Firms discount future cash flows at an interest rate of $r > 0$.

Consumers are uniformly distributed on the unit interval.[6] They have unit demands and incur transportation costs $b > 0$ per unit of distance. A consumer derives a surplus gross of transportation costs and price of $a > 0$ from consumption. Suppose that firm i is located at x_i and charges a price p_i. Then, the utility of a consumer located at z from buying from firm i is

$$a - b(z - x_i)^2 - p_i. \tag{1}$$

The[7] total mass of consumers at time t is $m(t)$, where $m' > 0$. We assume market size grows at most exponentially at a rate less than r to ensure that the NPV of future cash flows remains bounded.

Taking their locations as given, firms then compete in prices. The marginal cost of production is $c \geq 0$. We assume $\frac{a-c}{b} > 3$ to ensure the market is fully covered. That is, in equilibrium, each consumer prefers buying from either firm 1 or firm 2 over not buying. Let π^M denote a monopolist's instantaneous profits when market size is normalized to unity. Similarly, let π_i^D denote firm i's instantaneous profits when there are two firms in the market and the market size is normalized to unity. In online Appendix A, we show that

$$\pi^M(x) = \begin{cases} a - c - b(1-x)^2 & \text{if} \quad x \leq \frac{1}{2}, \\ a - c - bx^2 & \text{if} \quad x > \frac{1}{2}, \end{cases} \tag{2}$$

[5] This assumption is shared by Schmidt-Dengler (2006), Igami and Yang (2016), and Fang and Yang (2019, 2022).

[6] Tabuchi and Thisse (1995) and Anderson et al. (1997) consider general distributions.

[7] By contrast, Hotelling (1929) assumes that the consumer's utility is $a - b|z - x_i| - p_i$. As pointed out by d'Aspremont et al. (1979), this raises existence issues when firms compete in prices taking their locations as given. We avoid these issue by specifying quadratic transportation costs.

TABLE 1 Instantaneous Profits ($\pi_1^D(x_1, x_2)$, $\pi_2^D(x_1, x_2)$)

	$x_2 = \frac{1}{2}$	$x_2 = 1$
$x_1 = 0$	$(\frac{25b}{144}, \frac{49b}{144})$	$(\frac{b}{2}, \frac{b}{2})$
$x_1 = \frac{1}{2}$	(0,0)	$(\frac{49b}{144}, \frac{25b}{144})$

$$\pi_1^D(x_1, x_2) = \begin{cases} \frac{b(x_2-x_1)}{18}(2 + x_1 + x_2)^2 & \text{if} \quad x_1 \le x_2, \\ \frac{b(x_1-x_2)}{18}(4 - x_1 - x_2)^2 & \text{if} \quad x_1 > x_2, \end{cases} \tag{3}$$

$$\pi_2^D(x_1, x_2) = \begin{cases} \frac{b(x_2-x_1)}{18}(4 - x_1 - x_2)^2 & \text{if} \quad x_1 \le x_2, \\ \frac{b(x_1-x_2)}{18}(2 + x_1 + x_2)^2 & \text{if} \quad x_1 > x_2. \end{cases} \tag{4}$$

Note that $\pi^M(x)$ is maximal at $x = \frac{1}{2}$, whereas $\pi^D(x_1, x_2)$ is maximal at $x_1 = 0$ and $x_2 = 1$ as well as $x_1 = 1$ and $x_2 = 0$. Moreover, $\pi^M(x) \in \left[a - c - b, a - c - \frac{b}{4}\right]$ and $\pi_i^D(x_1, x_2) \in \left[0, \frac{b}{2}\right]$. Consequently, $\frac{a-c}{b} > 3$ implies $\pi^M(x) > \pi_i^D(x_1, x_2)$. Table 1 lists the instantaneous profits $(\pi_1^D(x_1, x_2), \pi_2^D(x_1, x_2))$ for selected combinations of x_1 and x_2.

The existing literature commonly assumes that firms play a static game of location choice followed by price competition. In these games, location choices depend on the number of entrants. The monopolist prefers to be located in the center ($x = \frac{1}{2}$) of the market rather than at its extremes ($x = 0$ or $x = 1$), because $\pi^M(x)$ is maximal at $x = \frac{1}{2}$. That is, the monopolist wants to be where demand is. For a duopolist, on the other hand, we have $\pi_1^D(x, x) = \pi_2^D(x, x) = 0$, as price competition drives profits down to zero if the firms' products are the same. Moreover, $\frac{\partial \pi_1^D}{\partial x_1} < 0$ and $\frac{\partial \pi_2^D}{\partial x_2} > 0$ if $x_1 < x_2$, whereas $\frac{\partial \pi_1^D}{\partial x_1} > 0$ and $\frac{\partial \pi_2^D}{\partial x_2} < 0$ if $x_1 > x_2$. Hence, each firm has an incentive to differentiate its product from that of its rival. In fact, firm *i*'s best reply to firm *j*'s location choice is given by

$$x_i^\circ(x_j) = \begin{cases} 0 & \text{if} \quad x_j > \frac{1}{2}, \\ \{0, 1\} & \text{if} \quad x_j = \frac{1}{2}, \\ 1 & \text{if} \quad x_j < \frac{1}{2}. \end{cases} \tag{5}$$

Consequently, duopolists prefer to locate at opposite ends of the market, and $x_i = 0$ and $x_j = 1$ is the outcome of a subgame perfect Nash equilibrium of the static game of (sequential or simultaneous) location choice followed by price competition. This is of course the principle of maximum differentiation, which stems from the fact that product differentiation alleviates price competition d'Aspremont et al. (1979).[8]

Rather than looking at a static game of location choice followed by price competition, we specify a fully dynamic entry process. Suppose firm 1 enters the market before firm 2. Because $t_1 \le t_2$, we call firm 1 the leader and firm 2 the follower. The NPV of the leader's payoffs is

$$V_1(t_1, t_2, x_1, x_2) = V_1^M(t_1, t_2, x_1, x_2) + V_1^D(t_1, t_2, x_1, x_2) - e^{-rt_1}F, \tag{6}$$

where

$$V_1^M(t_1, t_2, x_1, x_2) = \int_{t_1}^{t_2} e^{-rt}\pi^M(x_1)m(t)dt,$$

$$V_1^D(t_1, t_2, x_1, x_2) = \int_{t_2}^{\infty} e^{-rt}\pi_1^D(x_1, x_2)m(t)dt, \tag{7}$$

[8] Economides (1986) considers transportation costs of the form $b|z - x_i|^d$ and shows that maximum differentiation results if $d > 1.67$ but not if $1.26 < d < 1.67$. If $d < 1.26$, there does not exist an equilibrium in the two-stage game of location choice followed by price competition.

and the NPV of the follower's payoffs is

$$V_2(t_2, x_1, x_2) = \int_{t_2}^{\infty} e^{-rt}\pi_2^D(x_1, x_2)m(t)dt - e^{-rt_2}F. \tag{8}$$

Analogous expressions arise if firm 2 enters the market before firm 1.

Before turning to a characterization of the equilibrium, we impose two assumptions on the size of the market.

Assumption 1. $m(0) = 0$.

Assumption 2. There exists a $T < \infty$ such that $\pi_2^D(\frac{1}{2}, 1)m(T) > rF$.

Assumption 1 stipulates that the market takes off at time zero. Assumption 2 ensures that both firms eventually enter the market. To see this, note that $\pi_2^D(\frac{1}{2}, 1)$ is the minmax instantaneous profit of firm 2, which in turn is equal to the minmax instantaneous profit of firm 1. Hence, Assumption 2 requires that there exists a time T at which the market is so large that the NPV of the worst-case profits $\pi_2^D(\frac{1}{2}, 1)m(T)/r$ covers the fixed cost of entry F.

3. Sequential pre-commitment equilibrium

■ In SEQPE, at the outset of the game, each firm must irreversibly commit itself to a time and location at which it enters the market. These decisions are made sequentially. The leader therefore takes the follower's reactions into account when making its own decisions. In deriving the SEQPE, we assume that firm 1 enters before firm 2. Hence, given a SEQPE in which firm 1 is the leader and firm 2 the follower, there is a corresponding SEQPE in which firm 1 is the follower and firm 2 the leader.

The follower's problem. The follower solves

$$\max_{t_2 \geq 0, x_2 \in [0,1]} V_2(t_2, x_1, x_2). \tag{9}$$

The derivatives of V_2 with respect to t_2 and x_2 are

$$\frac{\partial V_2}{\partial t_2} = -e^{-rt_2}\pi_2^D(x_1, x_2)m(t_2) + re^{-rt_2}F, \tag{10}$$

$$\frac{\partial V_2}{\partial x_2} = \int_{t_2}^{\infty} e^{-rt}\frac{\partial \pi_2^D(x_1, x_2)}{\partial x_2}m(t)dt. \tag{11}$$

From this, it is clear that the follower's timing and location decisions depend solely on its profits from the duopoly phase.

Let $t_2^*(.)$ and $x_2^*(.)$ denote a solution to the follower's problem as a function of x_1.

Proposition 1. $t_2^*(x_1) = m^{-1}\left(\frac{rF}{\pi_2^D(x_1, x_2^*(x_1))}\right) \in (0, T)$ and $x_2^*(x_1) \in x_2^{\circ}(x_1)$, as defined in equation (5).

In what follows, we assume without loss of generality that $x_1 \in [0, \frac{1}{2}]$, which implies $\pi_2^D(x_1, 1) \geq \pi_2^D(x_1, 0)$. That is, $x_2^*(x_1) = 0$ is weakly dominated by $x_2^*(x_1) = 1$. To simplify the exposition, we thus set $x_2^*(x_1) = 1$ and $t_2^*(x_1) = m^{-1}\left(\frac{rF}{\pi_2^D(x_1, 1)}\right)$. Note that $t_2^{*\prime}(x_1) > 0$, because price competition intensifies as the leader moves toward the center of the market, which in turn reduces the follower's instantaneous profits. Hence, the follower has to wait longer until the size of the market has reached a level that allows for profitable entry.

The leader's problem. Taking the follower's reactions into account, the leader solves

$$\max_{t_1 \geq 0, x_1 \in [0, \frac{1}{2}]} V_1(t_1, t_2^*(x_1), x_1, 1). \tag{12}$$

The derivatives of V_1 with respect to t_1 and x_1 are

$$\frac{\partial V_1}{\partial t_1} = -e^{-rt_1}\pi^M(x_1)m(t_1) + re^{-rt_1}F, \tag{13}$$

$$\begin{aligned}\frac{\partial V_1}{\partial x_1} = & \int_{t_1}^{t_2^*(x_1)} e^{-rt}\pi^{M\prime}(x_1)m(t)dt \\ & + e^{-rt_2^*(x_1)}\pi^M(x_1)m(t_2^*(x_1))t_2^{*\prime}(x_1) \\ & + \int_{t_2^*(x_1)}^{\infty} e^{-rt}\frac{\partial \pi_1^D(x_1, 1)}{\partial x_1}m(t)dt \\ & - e^{-rt_2^*(x_1)}\pi_1^D(x_1, 1)m(t_2^*(x_1))t_2^{*\prime}(x_1).\end{aligned} \tag{14}$$

The leader's timing decision depends on its profits from the monopoly phase. Its location decision is governed by its profits from the monopoly phase and its profits from the duopoly phase. Specifically, the leader's location decision is governed by three considerations. First, by moving toward the center of the market, the leader increases its profits from the monopoly phase because $\pi^{M\prime} > 0$ (first term). Second, the leader decreases its profits from the duopoly phase because $\frac{\partial \pi_1^D}{\partial x_1} < 0$ (third term). Third, by moving toward the center of the market, the leader deters entry by the follower because $t_2^{*\prime} > 0$. This increases the duration of the monopoly phase (second term) and decreases the duration of the duopoly phase (fourth term). The net effect is positive because $\pi^M(x_1) > \pi_1^D(x_1, 1)$.

Let t_1^* and x_1^* denote a solution to the leader's problem.

Proposition 2. $t_1^* = m^{-1}\left(\frac{rF}{\pi^M(x_1^*)}\right) \in (0, T)$.

Unfortunately, we cannot say much about the leader's location decision due to the trade-off between the monopoly and the duopoly phases. Zhou and Vertinsky (2001) assume that the market grows linearly without bound and show that the leader chooses either the extreme ($x_1^* = 0$) or the central location ($x_1^* = \frac{1}{2}$) but never an intermediate one ($x_1^* \in (0, \frac{1}{2})$). That is, with a linear market size function, one phase is always much more "important" than the other. By contrast, we consider a general market size function.

SEQPE. The following proposition shows that, in any SEQPE, the payoff to the leader exceeds the payoff to the follower.[9]

Proposition 3. In any SEQPE, $V_1(t_1^*, t_2^*(x_1^*), x_1^*, 1) > V_2(t_2^*(x_1^*), x_1^*, 1)$.

Consequently, the follower has an incentive to become the leader, and it is not innocuous to pre-specify the order of entry. To see this, consider a SEQPE with outcome $(t_1^*, t_2^*(x_1^*), x_1^*, 1)$. Because the leader's payoff exceeds the follower's, firm 2 could imitate firm 1 by entering at time

[9] Proposition 2 allows us to replace the leader's problem with $\max_{t_1 \in [0,T], x_1 \in [0, \frac{1}{2}]} V_1(t_1, t_2^*(x_1), x_1, 1)$, where $t_2^*(x_1) = m^{-1}\left(\frac{rF}{\pi_2^D(x_1, 1)}\right)$ by Proposition 1. Because V_1 and $t_2^*(.)$ are continuous functions, a solution exists and therefore a SEQPE exists. Of course, there may be more than one SEQPE. Note also that $\pi^M(x_1^*) > \pi_2^D(x_1^*, 1)$ implies $t_1^* = m^{-1}\left(\frac{rF}{\pi^M(x_1^*)}\right) < m^{-1}\left(\frac{rF}{\pi_2^D(x_1^*, 1)}\right) = t_2^*(x_1^*)$. Hence, our assumption that firm 1 enters before firm 2 is warranted.

$t_1^* - \epsilon$, where $\epsilon > 0$, at location x_1^* and secure itself a payoff of $V_1(t_1^* - \epsilon, t_2^*(x_1^*), x_1^*, 1)$. Given that V_1 is continuous, the resulting payoff for firm 2 is arbitrarily close to firm 1's payoff in the SEQPE and thus in excess of firm 2's payoff in the SEQPE. Because the follower has an incentive to become the leader, the SEQPE is not subgame perfect Fudenberg and Tirole (1985).

4. Subgame perfect equilibrium

■ In a SEQPE, firms pre-commit to a time and location of entry, with preemption ruled out by assumption. We avoid this assumption here by specifying a fully dynamic entry process and characterizing the SPE of the dynamic game. In contrast to a SEQPE, the roles of leader and follower are determined endogenously in a SPE. Hence, the threat of preemption is operative and governs the times and locations of entry in a growing market, as we show below.

To avoid the technical difficulties associated with simultaneous actions in continuous-time games Simon and Stinchcombe (1989), we follow Gilbert and Harris (1984) and introduce the concept of decision lags. Let $I(t)$ denote the common information on the state of the game at time t. We assume that firm i with decision lag $h_i > 0$ can take an action at time t depending on the information available at time $t - h_i$, as given by $I(t - h_i)$. We further assume that $h_1 < h_2$, so that firm 1 is able to take an action at time t using more recent information than firm 2. We let the decision lags approach zero while maintaining $h_1 < h_2$. In the limit, the delay between information and action is negligible. Both firms observe $I(t)$ at time t, but the action taken by firm 1 is realized first and instantaneously incorporated into the information available to firm 2. Gilbert and Harris (1984) show that the first-mover advantage resulting from decision lags is trivial; the decision lags merely serve to rule out simultaneous entry when it is not optimal for both firms to enter at that time.

The follower's problem. Suppose firm 1 has just entered the market at time t_1 in location x_1 and firm 2 has not entered the market yet. That is, firm 1 is the leader and firm 2 the follower. Then, firm 2 solves

$$\max_{t_2 \geq t_1, x_2 \in [0,1]} V_2(t_2, x_1, x_2). \tag{15}$$

Note that we now require $t_2 \geq t_1$ as opposed to $t_2 \geq 0$ in Section 3, which reflects the dynamic nature of the entry process: Once time t_1 is reached, there is no going back, and the follower has to choose between entering now or in the future. Let $t_2^*(.)$ and $x_2^*(.)$ denote a solution to the follower's problem as a function of t_1 and x_1.

Proposition 4. Suppose $t_1 < \infty$. Then $t_2^*(t_1, x_1) = \max\left\{t_1, m^{-1}\left(\frac{rF}{\pi_2^D(x_1, x_2^*(x_1))}\right)\right\}$, where $m^{-1}\left(\frac{rF}{\pi_2^D(x_1, x_2^*(x_1))}\right) \in (0, T)$, and $x_2^*(x_1) \in x_2^\circ(x_1)$, as defined in equation (5).

In what follows, we assume without loss of generality that $x_1 \in [0, \frac{1}{2}]$, which implies $\pi_2^D(x_1, 1) \geq \pi_2^D(x_1, 0)$. That is, $x_2^*(x_1) = 0$ is weakly dominated by $x_2^*(x_1) = 1$. To simplify the exposition, we thus set $x_2^*(x_1) = 1$ and $t_2^*(t_1, x_1) = \max\left\{t_1, m^{-1}\left(\frac{rF}{\pi_2^D(x_1, 1)}\right)\right\}$. Note that $t_2^*(t_1, x_1)$ is nondecreasing in both of its arguments.

The leader's problem. Suppose firm 1 is about to enter the market at time t_1 and firm 2 has not entered the market yet. Taking the reactions of the lagging firm 2 into account, the leading firm 1 solves

$$\max_{x_1 \in [0, \frac{1}{2}]} V_1(t_1, t_2^*(t_1, x_1), x_1, 1). \tag{16}$$

Let $X_1^*(.)$ denote the set of solutions to the leader's problem as a correspondence of t_1. Because V_1 and $t_2^*(.)$ are continuous functions and $[0, \frac{1}{2}]$ is a compact set, the Theorem of the Maximum implies that $X_1^*(.)$ is nonempty and has a closed graph, and also that $V_1(t_1, t_2^*(t_1, x_1^*(t_1)), x_1^*(t_1), 1)$, where $x_1^*(t_1) \in X_1^*(t_1)$, is a continuous function of t_1.

To assess the properties of $X_1^*(.)$, fix t_1 and split $[0, \frac{1}{2}]$ into two (possibly empty) subsets $\underline{X}_1(t_1) = \left\{x_1 \in [0, \frac{1}{2}] \middle| t_1 \geq m^{-1}\left(\frac{rF}{\pi_2^D(x_1,1)}\right)\right\}$ and $\overline{X}_1(t_1) = \left\{x_1 \in [0, \frac{1}{2}] \middle| t_1 \leq m^{-1}\left(\frac{rF}{\pi_2^D(x_1,1)}\right)\right\}$. Hence, if the leader enters at time t_1 at any location $x_1 \in \underline{X}_1(t_1)$, then this triggers immediate entry by the follower, whereas any location $x_1 \in \overline{X}_1(t_1)$ leads to deferred entry. In what follows, we refer to immediate entry by the follower as case 1 and to deferred entry by the follower as case 2.

In online Appendix B, we characterize the two subsets $\underline{X}_1(t_1)$ and $\overline{X}_1(t_1)$ and the solution to the leader's problem on each of them. In case 1 (immediate entry by the follower), the leader maximizes instantaneous profits by locating at the extreme of the market. In case 2 (delayed entry by the follower), the leader's location decision is governed by similar considerations to those the leader faces in a SEQPE. In particular, by moving toward the center of the market, the leader increases its profits from the monopoly phase, deceases its profits from the duopoly phase, and increases the duration of the monopoly phase. Although the solution to the leader's problem on the set $\overline{X}_1(t_1)$ may not be unique, in Lemma 1 in online Appendix B, we provide conditions under which the set of solutions $\overline{X}_1^*(t_1)$ shifts to the left—that is, toward more extreme locations—with t_1.

Because $\pi_2^D(x_1, 1) \leq \pi_2^D(0, 1)$ and $\pi_2^D(x_1, 1) \geq \pi_2^D(\frac{1}{2}, 1)$, the solution to the leader's problem must be given by case 2 (deferred entry by the follower) if $t_1 \in \left[0, m^{-1}\left(\frac{rF}{\pi_2^D(0,1)}\right)\right) = \left[0, m^{-1}\left(\frac{2rF}{b}\right)\right)$ and by case 1 (immediate entry by the follower) if $t_1 \in \left(m^{-1}\left(\frac{rF}{\pi_2^D(\frac{1}{2},1)}\right), \infty\right) = \left(m^{-1}\left(\frac{144rF}{25b}\right), \infty\right)$, as can be seen in Table 1. If $t_1 \in \left[m^{-1}\left(\frac{2rF}{b}\right), m^{-1}\left(\frac{144rF}{25b}\right)\right]$, then the solution may be given by case 1 or case 2 or both. More formally, we therefore have

$$X_1^*(t_1) \subseteq \begin{cases} [0, \frac{1}{2}] & \text{if} \quad t_1 \in \left[0, m^{-1}\left(\frac{2rF}{b}\right)\right), \\ \{0\} \cup [\tilde{x}_1(t_1), \frac{1}{2}] & \text{if} \quad t_1 \in \left[m^{-1}\left(\frac{2rF}{b}\right), m^{-1}\left(\frac{144rF}{25b}\right)\right], \\ \{0\} & \text{if} \quad t_1 \in \left(m^{-1}\left(\frac{144rF}{25b}\right), \infty\right), \end{cases} \tag{17}$$

where $\tilde{x}_1(t_1)$ is the unique solution to $t_1 = m^{-1}\left(\frac{rF}{\pi_2^D(x_1,1)}\right)$, $\tilde{x}_1' > 0$, $\tilde{x}_1(t_1) = 0$ at $t_1 = m^{-1}\left(\frac{2rF}{b}\right)$, and $\tilde{x}_1(t_1) = \frac{1}{2}$ at $t_1 = m^{-1}\left(\frac{144rF}{25b}\right)$. This is in line with our intuition: If the leader enters early, the follower defers entry, irrespective of the leader's location. Knowing this, the leader faces a trade-off between the monopoly and the duopoly phases, and we are unable to pinpoint its location. By contrast, if the leader enters late, the follower enters immediately, irrespective of the leader's location. Knowing this, the leader chooses the extreme location. Finally, given intermediate times of entry, more extreme locations trigger intermediate entry, whereas more central locations lead to deferred entry by the follower.

Although we are in general unable to pinpoint the leader's location choice, we are able to determine how it changes over time. Consider first times of entry in the interval $\left[0, m^{-1}\left(\frac{2rF}{b}\right)\right)$. A straightforward implication of Lemma 1 in online Appendix B is that a late first entrant does not choose a less extreme location than an early first entrant. This is the content of the following proposition.

Proposition 5. Suppose $t_1, t_1' \in \left[0, m^{-1}\left(\frac{2rF}{b}\right)\right)$ and $t_1 < t_1'$. Then $\min X_1^*(t_1) \geq \max X_1^*(t_1')$ with strict inequality if $\min X_1^*(t_1) \in \left(0, \frac{1}{2}\right)$.

Corollary 1 summarizes the implications of Proposition 5 for the follower's payoff.

Corollary 1. Suppose $t_1, t_1' \in \left[0, m^{-1}\left(\frac{2rF}{b}\right)\right)$ and $t_1 < t_1'$. Then, $\max_{x_1 \in X_1^*(t_1)} V_2(t_2^*(t_1, x_1), x_1, 1) \leq \min_{x_1' \in X_1^*(t_1')} V_2(t_2^*(t_1', x_1'), x_1', 1)$ with strict inequality if $\min X_1^*(t_1) \in \left(0, \frac{1}{2}\right)$.

Consider next the times of entry in the interval $\left[m^{-1}\left(\frac{2rF}{b}\right), m^{-1}\left(\frac{144rF}{25b}\right)\right]$. Once the leader chooses to locate at the extreme of the market, it continues to do so, which clarifies the relationship between cases 1 and 2. Because we do not use this fact in our characterization of the SPE, we omit the argument.

SPE. Define $L(t_1)$ to be the payoff of the firm that enters first at time t_1, thereby preempting its rival, and define $F(t_1)$ to be the payoff of the firm that enters second and is thus being preempted by its rival at time t_1:

$$L(t_1) = V_1(t_1, t_2^*(t_1, x_1^*(t_1)), x_1^*(t_1), 1),$$
$$F(t_1) = V_2(t_2^*(t_1, x_1^*(t_1)), x_1^*(t_1), 1),$$

where $x_1^*(t_1) \in X_1^*(t_1)$. We sometimes write $L(t_1; x_1^*(t_1))$ and $F(t_1; x_1^*(t_1))$ to show explicitly how payoffs depend on time t_1 and location $x_1^*(t_1)$. Recall that $L(.)$ is a continuous function of t_1. By contrast, $F(.)$ is a correspondence of t_1 because $X_1^*(t_1)$, the set of solutions to the leader's problem at t_1, is not guaranteed to be a singleton. The correspondence $F(.)$ is nonempty and has a closed graph. To simplify the exposition, we write $F(t_1) \geq y$ as shorthand for $\min_{x_1 \in X_1^*(t_1)} F(t_1; x_1) \geq y$, $F(t_1) < y$ for $\max_{x_1 \in X_1^*(t_1)} F(t_1; x_1) < y$, and so on.

Consider the subgame starting at time t_1. There are two classes of subgames, namely the ones where some firm has already entered the market and the ones where no firm has entered the market yet. The first class is easy to deal with because the problem of the remaining firm boils down to the follower's problem that we analyzed above. We therefore focus on the second class in what follows.

To characterize the SPE, we partition the time axis as follows:

- *Region 1*: $\left[0, m^{-1}\left(\frac{rF}{a-c-\frac{b}{4}}\right)\right)$,
- *Region 2*: $\left[m^{-1}\left(\frac{rF}{a-c-\frac{b}{4}}\right), m^{-1}\left(\frac{rF}{a-c-b}\right)\right]$,
- *Region 3*: $\left(m^{-1}\left(\frac{rF}{a-c-b}\right), m^{-1}\left(\frac{2rF}{b}\right)\right)$,
- *Region 4*: $\left[m^{-1}\left(\frac{2rF}{b}\right), m^{-1}\left(\frac{144rF}{25b}\right)\right]$,
- *Region 5*: $\left(m^{-1}\left(\frac{144rF}{25b}\right), \infty\right)$.

Recall that the solution to the leader's problem is given by case 2 (deferred entry by the follower) in regions 1, 2, and 3 and by case 1 (immediate entry by the follower) in region 5. In region 4, the solution may be given by case 1 or case 2 or both. Recall further that $F(.)$ is nondecreasing in regions 1, 2, and 3 by Corollary 1.

In online Appendix C, we show that $L(.)$ is increasing in region 1 and decreasing in regions 3, 4, and 5. Hence, $L(.)$ attains a global maximum in region 2. In what follows, we assume that the global maximum of $L(.)$ is unique and that $L(.)$ is increasing to the left of the global maximum and decreasing to the right.

Assumption 3. $L(.)$ is unimodal.

A sufficient condition for Assumption 3 to hold is that there exists x_1 such that $x_1 \in X_1^*(t_1)$ for all $t_1 \in \left[m^{-1}\left(\frac{rF}{a-c-\frac{b}{4}}\right), m^{-1}\left(\frac{rF}{a-c-b}\right)\right]$. The reason is that, holding x_1 fixed, $L(t_1) = V_1\left(t_1, m^{-1}\left(\frac{rF}{\pi_2^D(x_1,1)}\right), x_1, 1\right)$ is strictly quasiconcave in t_1. Moreover, Assumption 3 always held in the numerical examples we studied in Section 6.

Define $t_1^* = \arg\max_{t_1 \geq 0} L(t_1)$ and further partition region 2 by the global maximum of $L(.)$ as follows:

- *Region 2a*: $\left[m^{-1}\left(\frac{rF}{a-c-\frac{b}{4}}\right), t_1^*\right)$,
- *Region 2b*: $\left[t_1^*, m^{-1}\left(\frac{rF}{a-c-b}\right)\right]$.

Regions 2b, 3, 4, and 5. Working backward through time, consider the subgame starting at time $t_1 \in \left[t_1^*, \infty\right)$. In online Appendix C, we show that $L(t_1) \geq F(t_1)$ and

$$L(t_1) > \max\left\{L(t_1'), F(t_1')\right\} \tag{18}$$

for all $t_1' > t_1$. Because the payoff to entering first and becoming the leader at time t_1 is at least as large as the payoff to becoming the follower at time t_1 and larger than the payoff to becoming either the leader or the follower at some later time t_1', it is optimal for a firm to enter at time t_1 and location $x_1 \in X_1^*(t_1)$. Moreover, it is optimal to enter irrespective of the opponent's strategy. Hence, firm 1 enters as it has a shorter decision lag than firm 2. Note that to the extent that $X_1^*(t_1)$ is not a singleton, there may be multiplicity, though this multiplicity has no impact on the outcome of the SPE because, as we show below, the time of first entry is prior to t_1.

Regions 1 and 2a. Consider times of entry in the interval $[0, t_1^*)$. Rent equalization occurs if $F(.)$ cuts $L(.)$ from above at some t_1 to the left of the global maximum of $L(.)$. In what follows, we assume that $F(0) > L(0)$. In essence, this requires that the fixed cost of entry F is sufficiently large.[10] The following proposition shows that $L(t_1^*) > F(t_1^*)$. Note that it does not presume that the global maximum of $L(.)$ is unique.

Proposition 6. Let $t_1^* \in \arg\max_{t_1 \geq 0} L(t_1)$. Then, $L(t_1^*) > F(t_1^*)$.

It turns out that there may be more than one SPE. Before stating our main theorem, we set up some notation that allows us to describe the set of SPEs. Define

$$\hat{t}_1 = \sup\left\{t_1 \in [0, t_1^*] | F(t_1) \geq L(t_1)\right\}, \tag{19}$$

where $F(t_1) \geq L(t_1)$ is once again shorthand for $\min_{x_1 \in X_1^*(t_1)} F(t_1; x_1) \geq L(t_1)$, and

$$\tilde{T}_1 = \left\{ t_1 \in (\hat{t}_1, t_1^*] \,\middle|\, \max_{x_1 \in X_1^*(t_1)} F(t_1; x_1) \geq L(t_1) \right\}. \tag{20}$$

The sup here accounts for the possibility that $F(.)$ coincides with $L(.)$ up to (but not including) some point. The upper bound t_1^* is necessary because we know that $F(.)$ eventually coincides with $L(.)$.

We know that $\hat{t}_1 \in (0, t_1^*)$ exists because $F(0) > L(0)$ and $F(t_1^*) < L(t_1^*)$ by Proposition 6. By contrast, $\tilde{T}_1 \subseteq (\hat{t}_1, t_1^*)$ may be empty. The following proposition summarizes the properties of $\hat{t}_1$ and $\tilde{T}_1$.

Proposition 7. (i) $X_1^*(\hat{t}_1)$ is a singleton; (ii) $F(\hat{t}_1) = L(\hat{t}_1)$; (iii) $X_1^*(\tilde{t}_1)$ is not a singleton for any $\tilde{t}_1 \in \tilde{T}_1$; (iv) $\max_{x_1 \in X_1^*(\tilde{t}_1)} F(\tilde{t}_1; x_1) = L(\tilde{t}_1)$ for all $\tilde{t}_1 \in \tilde{T}_1$; (v) $\tilde{T}_1$ consists of isolated points.

To illustrate Proposition 7, suppose $\tilde{T}_1 = \varnothing$. Then $F(.)$ cuts $L(.)$ from above at $\hat{t}_1$ and stays below $L(.)$ to the right of $\hat{t}_1$. By contrast, if $\tilde{T}_1 \neq \varnothing$, then $F(.)$ touches $L(.)$ from below at $\tilde{t}_1 \in \tilde{T}_1$ to the right of $\hat{t}_1$. In this sense, $\tilde{T}_1 \neq \varnothing$ is a knife-edge case.

[10] This assumption is made purely for analytical convenience. In fact, as long as we are willing to extend the time axis below zero (as we do in Section 6), there always exists a $t_1 < t_1^*$ such that $F(t_1) > L(t_1)$. The reason is that $F(.)$ is positive by Assumption 2, whereas $\lim_{t_1 \to -\infty} L(t_1) = -\infty$.

This leads to our main theorem.

Theorem 1. (i) The following strategy is a SPE: If $t_1 < \hat{t}_1$, do not enter; if $t_1 \geq \hat{t}_1$ and no firm has entered the market yet, enter at location $\max X_1^*(t_1)$. (ii) For any $\tilde{t}_1 \in \tilde{T}_1$, the following strategy is a SPE: If $t_1 < \tilde{t}_1$, do not enter; if $t_1 \geq \tilde{t}$ and no firm has entered the market yet, enter at location $\min X_1^*(t_1)$.

Recall that we focus on subgames where no firm has entered the market yet. To relate Theorem 1 to equations (19) and (20), note that $\min_{x_1 \in X_1^*(t_1)} F(t_1; x_1) = F(t_1; \max X_1^*(t_1))$ and $\max_{x_1 \in X_1^*(t_1)} F(t_1; x_1) = F(t_1; \min X_1^*(t_1))$ because $\frac{\partial \pi_2^D}{\partial x_1} < 0$ if $x_1 \leq x_2$.

We prove part (i) of Theorem 1 and relegate the proof of part (ii) to online Appendix D. Because we have already dealt with regions 2b, 3, 4, and 5, we focus on regions 1 and 2a in what follows. Working backward through time, consider the subgame starting at time $t_1 \in \left[\hat{t}_1, t_1^*\right]$. If a firm enters first at time t_1 and location $\max X_1^*(t_1)$ according to the prescribed strategy, then it gets $L(t_1)$. If the firm deviates from the prescribed strategy, then its rival enters first and the firm gets $F(t_1; \max X_1^*(t_1))$. We have $\min_{x_1 \in X_1^*(t_1)} F(t_1; x_1) \leq L(t_1)$ for all $t_1 \geq \hat{t}_1$ with strict inequality whenever $t_1 > \hat{t}_1$ by construction (see equation (19)) and $\min_{x_1 \in X_1^*(t_1)} F(t_1; x_1) = F(t_1; \max X_1^*(t_1))$. Hence, the firm has no incentive to deviate from the prescribed strategy. Because it has a shorter decision lag than firm 2, firm 1 enters.

Continuing to work backward through time, consider the subgame starting at time $t_1 \in \left[0, \hat{t}_1\right)$. If a firm does not enter according to the prescribed strategy, then it gets either $L(\hat{t}_1)$ or $F(\hat{t}_1; \max X_1^*(\hat{t}_1))$. If the firm deviates from the prescribed strategy and enters first at time t_1 at location $x_1 \in [0, \frac{1}{2}]$, then it gets at most $L(t_1)$ (and $L(t_1)$ if $x_1 \in X_1^*(t_1)$). We have $\min_{x_1 \in X_1^*(\hat{t}_1)} F(\hat{t}_1; x_1) \geq L(\hat{t}_1)$ by construction and $\min_{x_1 \in X_1^*(\hat{t}_1)} F(\hat{t}_1; x_1) = F(\hat{t}_1; \max X_1^*(\hat{t}_1))$. Moreover, $L(t_1) < L(\hat{t}_1)$ for all $t_1 < \hat{t}_1$ because $L(.)$ is increasing in region 1 and in region 2a by Assumption 3. Taken together, we have

$$L(t_1) < L(\hat{t}_1) \leq F(\hat{t}_1; \max X_1^*(\hat{t}_1)). \tag{21}$$

Hence, the firm has no incentive to deviate from the prescribed strategy.

5. Welfare

■ The nature of the game being played has stark welfare implications. Consider an omnipotent social planner who controls the time and location of entry as well as firms' pricing decisions in the product market. The planner's goal is to maximize social surplus, consisting of gains from trade net of transportation costs and fixed costs. Let ω^M and ω^D denote instantaneous social surplus under a monopoly and a duopoly, respectively. We have

$$\omega^M(x) = a - c - b\left(x^2 - x + \frac{1}{3}\right), \tag{22}$$

$$\omega^D(x_1, x_2) = \begin{cases} a - c - \frac{b}{3} + \frac{b(x_2 - x_1)}{4}(x_1 + x_2)^2 + bx_2(1 - x_2) & \text{if} \quad x_1 \leq x_2, \\ a - c - \frac{b}{3} + \frac{b(x_1 - x_2)}{4}(x_1 + x_2)^2 + bx_1(1 - x_1) & \text{if} \quad x_1 > x_2. \end{cases} \tag{23}$$

These expressions are independent of prices because demand is inelastic; they are derived in online Appendix A. Note that $\omega^M(x)$ is maximal at $x = \frac{1}{2}$ and $\omega^D(x_1, x_2)$ is maximal at $x_1 = \frac{1}{4}$ and $x_2 = \frac{3}{4}$ as well as $x_1 = \frac{3}{4}$ and $x_2 = \frac{1}{4}$. Moreover, $\omega^M(x) \in \left[a - c - \frac{b}{3}, a - c - \frac{b}{12}\right]$ and $\omega^D(x_1, x_2) \in \left[a - c - \frac{b}{3}, a - c - \frac{b}{48}\right]$. Two firms are therefore not necessarily better than one firm.

The NPV of social surplus is

$$W(t_1, t_2, x_1, x_2) = \int_{t_1}^{t_2} e^{-rt}\omega^M(x_1)m(t)dt + \int_{t_2}^{\infty} e^{-rt}\omega^D(x_1, x_2)m(t)dt - e^{-rt_1}F - e^{-rt_2}F. \tag{24}$$

We search for the times and locations of entry that maximize social surplus. That is, we solve the following problem:

$$\max_{0\le t_1\le t_2, 0\le x_1\le x_2\le 1, x_1\in[0,\frac{1}{2}]} W(t_1, t_2, x_1, x_2). \tag{25}$$

The planner may decide to have one instead of two firms in the market. In this case, the solution to the above problem is given by $t_1 = m^{-1}\left(\frac{rF}{\omega^M(\frac{1}{2})}\right)$ and $x_1 = \frac{1}{2}$ (along with $t_2 = \infty$ and $x_2 \in [\frac{1}{2}, 1]$).

6. Numerical analysis

■ We use numerical analysis to contrast the outcomes of the game under SEQPE and SPE and to illustrate their welfare implications. To model a wide range of possible market size functions, we choose the piecewise linear specification

$$m(t) = \begin{cases} 0 & \text{if} \quad t \le 0, \\ \frac{\mu_1}{\tau_1}t & \text{if} \quad 0 < t \le \tau_1, \\ \mu_1 + \frac{1-\mu_1}{\tau_2-\tau_1}(t-\tau_1) & \text{if} \quad \tau_1 < t \le \tau_2, \\ 1 & \text{if} \quad t > \tau_2, \end{cases} \tag{26}$$

where $\tau_1 > 0$ and $\tau_2 > \tau_1$ denote the end of the first and second growth periods, respectively, and $0 < \mu_1 < 1$ is the market size at the end of the first growth period. Note that market size is bounded. Moreover, $m(.)$ is concave, linear, or convex in $t \in [0, \tau_2]$ depending on whether $\frac{\mu_1}{\tau_1} \gtreqless \frac{1-\mu_1}{\tau_2-\tau_1}$.

The structure of the payoffs allows us to normalize some parameters. First, consumers' gross surplus a and firms' marginal cost c appear exclusively in the monopolist's profit function as $a - c$. Hence, without loss of generality we set $c = 0$. Second, if a, b, and F are multiplied by some nonzero constant, then this does not affect the follower's timing decision or the leader's location decision. Hence, $L(.)$ and $F(.)$ are homogeneous of degree one in the triple (a, b, F). We therefore normalize $b = 1$. Third, the interest rate r merely determines the time scale and is, therefore, not of interest by itself. We set $r = 0.05$ in what follows.

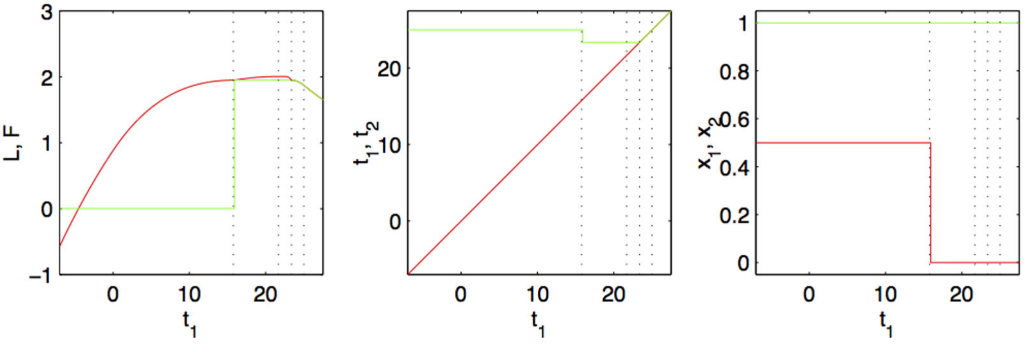
We conduct a search over the 15,000 parameterizations given by all possible combinations of $a \in \{3, 6, 9\}$, 10 equidistant values of $F \in \left(0, \frac{\pi_2^D(\frac{1}{2},1)}{r}\right)$ (see Assumption 2), $\tau_2 \in \{5, 10, 15, 25, 50\}$, 10 equidistant values of $\mu_1 \in (0, 1)$, and 10 equidistant values of $\tau_1 \in (0, \tau_2)$. That is, we choose τ_1 given τ_2 such that $\tau_1 < \tau_2$. Finally, as discussed in Section 4, we extend the time axis below zero. To compute $L(.)$ and $F(.)$, we discretize time with a period length of $\Delta = 0.01$.

For each of the parameterizations given above, our computations indicate that $L(.)$ is increasing to the right of t_1^* and decreasing to the left, thus justifying the simplifying assumption made in Theorem 1 in Section 4. Unimodality of $L(.)$ in turn ensures that the SEQPE is unique.[11] Moreover, our computations lead to a unique SPE in which $F(.)$ cuts $L(.)$ from above at $\hat{t}_1$ and stays below $L(.)$ to the right of $\hat{t}_1$ in accordance with part (i) of Theorem 1.

[11] Recall from Section 3 that, in the SEQPE, the leader enters at time t_1^*.

FIGURE 1

EXAMPLE OF A “TOUCHING” EQUILIBRIUM. PAYOFFS (LEFT PANEL), TIMES OF ENTRY (MIDDLE PANEL), AND LOCATIONS OF ENTRY (RIGHT PANEL)



A dark (bright) line designates the first (second) entrant. Parameters are $a = 3$, $b = 1$, $c = 0$, $F = 3.47$, $r = 0.05$, $\tau_1 = 22.73$, $\mu_1 = 0.09$, and $\tau_2 = 25$.

TABLE 2 Parameterizations with $t_1^{SPE} > t_1^{W}$

#	a	F	τ_1	μ_1	τ_2		t_1	t_2	x_1	L	F	C
1	3	3.16	9.09	0.09	10	SPE	5.43	9.32	0.00	4.22	4.22	23.92
						W	5.41	∞	0.50			
2	3	3.16	13.64	0.09	15	SPE	8.27	13.97	0.00	3.32	3.32	18.90
						W	8.12	∞	0.50			
3	3	3.16	22.73	0.09	25	SPE	14.15	23.29	0.00	2.05	2.05	11.81
						W	13.53	∞	0.50			

Note: $x_2^{SPE} = 1$ omitted.

Although our 15,000 parameterizations do not contain such a case, it is possible that $\tilde{T}_1 \neq \varnothing$ in accordance with part (ii) of Theorem 1. Figure 1 illustrates that there may exist one or more “touching” equilibria besides the “cutting” equilibrium. As can be seen, $F(.)$ cuts $L(.)$ from above at $\hat{t}_1 = -4.52$ and $F(.)$ touches $L(.)$ from below at $\tilde{t}_1 = 15.92$. Note, however, that $\tilde{T}_1 \neq \varnothing$ is a knife-edge case in the sense that, if the fixed cost of entry F is chosen to be slightly larger than in the example, $F(.)$ is below $L(.)$ but does not touch it, whereas $F(.)$ is above $L(.)$ if F is chosen to be slightly smaller.

In what follows, we use superscripts to distinguish between the outcome of the SEQPE, the outcome of the SPE, and the times and locations of entry chosen by the omnipotent social planner, indicated with a W superscript. Our first result summarizes the implications of preemption for the times of entry.

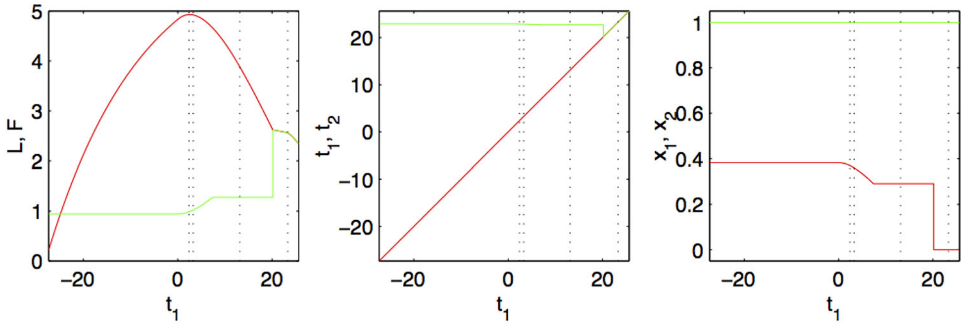
Result 1 (Times of entry). We have $t_1^W < t_1^{SEQPE}$ and $t_2^{SEQPE} \leq t_2^{SPE} < t_2^W$. Moreover, in 99.98% of parameterizations, we have $t_1^{SPE} < t_1^W$.

Note that $t_1^{SPE} < t_1^{SEQPE}$ by construction of the SPE. Proposition 5 then implies $x_1^{SEQPE} \leq x_1^{SPE}$ (as stated below in Result 2), and hence $t_2^{SEQPE} \leq t_2^{SPE}$. Table 2 lists the parameterizations with $t_1^{SPE} > t_1^W$, which we describe in greater detail below.

Result 1 says that, from a welfare point of view, the follower enters too early in a SPE as well as in a SEQPE. By contrast, the leader usually enters too early in the SPE but always too late in a SEQPE. In fact, we have $t_1^{SPE} < 0$ in 99.56% of parameterizations. This demonstrates the power of preemption: Because the incentive to preempt persists until payoffs are equalized,

FIGURE 2

EXAMPLE OF AN INTERMEDIATE LOCATION



Payoffs (left panel), times of entry (middle panel), and locations of entry (right panel). A dark (bright) line designates the first (second) entrant. Parameters are $a = 3$, $b = 1$, $c = 0$, $F = 1.58$, $r = 0.05$, $\tau_1 = 22.73$, $\mu_1 = 0.27$, and $\tau_2 = 25$.

the leader is forced to enter too early in a SPE, usually at a time at which there is not even any demand yet.

Our second result concerns the locations of entry. The result is in fact analytic, and we report it here merely for the sake of completeness.

Result 2 (Locations of entry). We have $x_1^{SEQPE} \leq x_1^{SPE}$ and $x_2^W \leq x_2^{SEQPE} = x_2^{SPE} = 1$.

By contrast, the relationship between x_1^W, x_1^{SEQPE}, and x_1^{SPE} is ambiguous. Result 2 says that, because $x_1^{SEQPE} \leq x_1^{SPE}$, a SPE entails less extreme locations and thus less product differentiation than a SEQPE.

Figure 2 presents an example. In the SPE, the first entry occurs at time $t_1^{SPE} = -24.86$ and location $x_1^{SPE} = 0.38$ and the second at time $t_2^{SPE} = 22.93$ and location $x_2^{SPE} = 1$; in the SEQPE, the first entry occurs at time $t_1^{SEQPE} = 2.53$ and location $x_1^{SEQPE} = 0.37$ and the second at time $t_2^{SEQPE} = 22.89$ and location $x_2^{SEQPE} = 1$. Note that this runs counter to Zhou and Vertinsky's (2001) claim that, in a SEQPE, the first entrant chooses either the extreme ($x_1 = 0$) or the central ($x_1 = \frac{1}{2}$) but never an intermediate ($x_1 \in \left(0, \frac{1}{2}\right)$) location. Of course, our model is more general than theirs because Zhou and Vertinsky (2001) assume that market size grows linearly without bound.

Given that both firms enter too early in a SPE but not in a SEQPE, one expects that the combined payoffs to firms and consumers are lower in a SPE than in a SEQPE. In stating the next result, we use L and F as shorthand for the payoff to the leader and the follower, respectively, and C as shorthand for the payoff to consumers. Expressions for instantaneous consumer surplus under monopoly and duopoly are derived in online Appendix A; the NPV of consumer surplus C is defined in the obvious way.

Result 3 (Welfare comparison). In 99.89% of parameterizations, we have $\Sigma^{SPE} \equiv L^{SPE} + F^{SPE} + C^{SPE} < L^{SEQPE} + F^{SEQPE} + C^{SEQPE} \equiv \Sigma^{SEQPE}$.

Table 3 lists the parameterizations with $\Sigma^{SPE} > \Sigma^{SEQPE}$. These 16 parameterizations have in common the fact that the market undergoes a long period of sluggish growth followed by a short period of rapid growth as τ_1 is large relative to τ_2 and μ_1 is small. Moreover, gross surplus a is low and the fixed cost of entry F is high. Note that the 3 parameterizations with $t_1^{SPE} > t_1^W$ in Table 2 reappear in the 16 parameterizations with $\Sigma^{SPE} > \Sigma^{SEQPE}$.

The following result quantifies the extent to which rents are dissipated under the two equilibrium concepts. We use W^W as shorthand for $W(t_1^W, t_2^W, x_1^W, x_2^W)$.

TABLE 3 Parameterizations with $\Sigma^{SPE} > \Sigma^{SEQPE}$

#	a	F	τ_1	μ_1	τ_2		t_1	t_2	x_1	L	F	C	Σ
1	3	2.53	9.09	0.09	10	SPE	3.30	9.25	0.00	4.62	4.62	23.99	33.23
						SEQPE	6.31	9.25	0.00	4.69	4.62	23.91	33.22
2	3	2.53	13.64	0.09	15	SPE	5.16	13.88	0.00	3.63	3.63	18.99	26.25
						SEQPE	9.47	13.88	0.00	3.72	3.63	18.89	26.24
3	3	2.53	22.73	0.09	25	SPE	9.23	23.13	0.00	2.25	2.24	11.92	16.41
						SEQPE	15.78	23.13	0.00	2.34	2.24	11.80	16.39
4	3	2.84	4.55	0.09	5	SPE	2.17	4.64	0.00	5.63	5.63	30.30	41.55
						SEQPE	3.55	4.64	0.00	5.66	5.63	30.25	41.54
5	3	2.84	8.18	0.09	10	SPE	3.34	8.57	0.00	4.50	4.50	24.54	33.54
						SEQPE	6.39	8.57	0.00	4.58	4.50	24.45	33.54
6	3	2.84	9.09	0.09	10	SPE	4.44	9.28	0.00	4.42	4.42	23.96	32.80
						SEQPE	7.10	9.28	0.00	4.48	4.42	23.88	32.78
7	3	2.84	12.27	0.09	15	SPE	5.21	12.85	0.00	3.57	3.57	19.64	26.78
						SEQPE	9.59	12.85	0.00	3.67	3.57	19.53	26.78
8	3	2.84	13.64	0.09	15	SPE	6.82	13.93	0.00	3.47	3.47	18.95	25.90
						SEQPE	10.65	13.93	0.00	3.54	3.47	18.85	25.86
9	3	2.84	22.73	0.09	25	SPE	11.86	23.21	0.00	2.15	2.15	11.86	16.16
						SEQPE	17.76	23.21	0.00	2.22	2.15	11.75	16.12
10	3	3.16	4.09	0.09	5	SPE	1.95	4.32	0.00	5.42	5.42	30.64	41.49
						SEQPE	3.55	4.32	0.00	5.47	5.42	30.59	41.48
11	3	3.16	4.55	0.09	5	SPE	2.68	4.66	0.00	5.38	5.38	30.28	41.03
						SEQPE	3.95	4.66	0.00	5.40	5.38	30.23	41.01
12	3	3.16	8.18	0.09	10	SPE	4.02	8.63	0.00	4.30	4.30	24.50	33.10
						SEQPE	7.10	8.63	0.00	4.38	4.30	24.41	33.08
13	3	3.16	9.09	0.09	10	SPE	5.43	9.32	0.00	4.22	4.22	23.92	32.37
						SEQPE	7.89	9.32	0.00	4.27	4.22	23.84	32.33
14	3	3.16	12.27	0.09	15	SPE	6.21	12.95	0.00	3.40	3.40	19.60	26.41
						SEQPE	10.65	12.95	0.00	3.50	3.40	19.48	26.39
15	3	3.16	13.64	0.09	15	SPE	8.27	13.97	0.00	3.32	3.32	18.90	25.54
						SEQPE	11.84	13.97	0.00	3.37	3.32	18.81	25.50
16	3	3.16	22.73	0.09	25	SPE	14.15	23.29	0.00	2.05	2.05	11.81	15.91
						SEQPE	19.73	23.29	0.00	2.10	2.05	11.71	15.86

Note: $x_2^{SPE} = 1$ and $x_2^{SEQPE} = 1$ omitted.

Result 4 (Rent dissipation). We have

$$0.21 < \frac{\Sigma^{SPE}}{W^W} < 1, \tag{27}$$

$$0.93 < \frac{\Sigma^{SEQPE}}{W^W} < 1, \tag{28}$$

$$0.95 < \frac{W^W - \Sigma^{SPE}}{W^W - \Sigma^{SEQPE}} < 237. \tag{29}$$

Hence, up to 79% of rents are dissipated by premature entry in a SPE as opposed to at most 7% in a SEQPE. This again demonstrates how powerful the incentive to preempt can be. Comparing rent dissipation under the two equilibrium concepts, Result 4 says that, at least for some parameterizations, 237 times as much surplus is wasted under the SPE than under the SEQPE. Tables 4–6 illustrate Result 4. Table 4 shows the parameterizations with the lowest value of $\frac{\Sigma^{SPE}}{W^W}$, Table 5 the ones with the lowest value of $\frac{\Sigma^{SEQPE}}{W^W}$, and Table 6 the ones with the highest value of $\frac{W^W - \Sigma^{SPE}}{W^W - \Sigma^{SEQPE}}$.

TABLE 4 Examples of Rent Dissipation: Lowest Values of $\frac{\Sigma^{SPE}}{W^W}$

#	a	F	τ_1	μ_1	τ_2		t_1	t_2	x_1	L	F	C	Σ, W	$\frac{\Sigma^{SPE}}{W^W}$
1	9	3.16	4.55	0.55	50	SPE	−66.78	40.91	0.50	0.07	0.03	23.07	23.17	0.21
						W	0.15	∞	0.50				109.35	
2	9	3.16	4.55	0.45	50	SPE	−64.96	42.42	0.50	0.05	0.03	21.42	21.50	0.22
						W	0.18	∞	0.50				99.95	
3	9	3.16	4.55	0.64	50	SPE	−68.24	38.64	0.50	0.06	0.03	25.68	25.77	0.22
						W	0.13	∞	0.50				118.76	
4	6	3.16	4.55	0.55	50	SPE	−58.48	40.91	0.50	0.05	0.03	15.45	15.53	0.22
						W	0.22	∞	0.50				71.51	
5	6	3.16	4.55	0.45	50	SPE	−56.66	42.42	0.50	0.05	0.03	14.33	14.41	0.22
						W	0.27	∞	0.50				65.27	

Note: $x_2^{SPE} = 1$ omitted.

TABLE 5 Examples of Rent Dissipation: Lowest Values of $\frac{\Sigma^{SEQPE}}{W^W}$

#	a	F	τ_1	μ_1	τ_2		t_1	t_2	x_1	L	F	C	Σ, W	$\frac{\Sigma^{SEQPE}}{W^W}$
1	3	3.16	22.73	0.09	25	SEQPE	19.73	23.29	0.00	2.10	2.05	11.71	15.86	0.93
						W	13.53	∞	0.50				16.97	
2	3	3.16	13.64	0.18	15	SEQPE	5.92	13.86	0.00	3.82	3.33	19.23	26.37	0.94
						W	4.06	∞	0.50				28.16	
3	3	3.16	12.27	0.09	15	SEQPE	10.65	12.95	0.00	3.50	3.40	19.48	26.39	0.94
						W	7.31	∞	0.50				28.16	
4	3	3.16	8.18	0.27	10	SEQPE	2.37	8.29	0.00	5.19	4.33	25.13	34.66	0.94
						W	1.62	∞	0.50				36.96	
5	3	3.16	9.09	0.27	10	SEQPE	2.63	9.14	0.00	5.14	4.24	24.47	33.86	0.94
						W	1.80	∞	0.50				36.09	

Note: $x_2^{SEQPE} = 1$ Omitted.

7. Empirical application

■ The nature of the game being played generates testable predictions. In our empirical application, we focus on the relationship between market size at the time of first entry and the market growth rate. Under a SEQPE, by Proposition 2 the threshold for market size at the time of first entry is

$$m(t_1^{SEQPE}) = \frac{rF}{\pi_1^M(x_1^{SEQPE})}. \tag{30}$$

Hence, conditional on the location of first entry, the market growth rate does not affect the threshold for market size at the time of first entry. This is not the case under SPE, however, where the threat of preemption is operative until rents equalize. Intuitively, as the market growth rate increases, the NPV of the leader's payoff increases more than the NPV of the follower's payoff. Rent equalization therefore requires earlier first entry.

To make this intuition concrete, we specify the exponential market size function $m(t) = e^{\gamma t}$, where $0 < \gamma < r$ is the rate of market growth. In what follows, we index the outcome of the game under SPE by γ and restrict attention to the empirically relevant "cutting" equilibria in part (i) of Theorem 1. Proposition 8 provides a prediction about the relationship between market size at the time of first entry and the rate of market growth under SPE that we use to test whether SEQPE or SPE better fits the data in a particular setting.

TABLE 6 Examples of Rent Dissipation: Highest Values of $\frac{W^W-\Sigma^{SPE}}{W^W-\Sigma^{SEQPE}}$

#	a	F	τ_1	μ_1	τ_2		t_1	t_2	x_1	L	F	C	Σ, W	$\frac{W^W-\Sigma^{SPE}}{W^W-\Sigma^{SEQPE}}$
1	9	3.16	9.09	0.55	50	SPE	−64.07	41.82	0.50	0.04	0.03	21.94	22.01	215.40
						SEQPE	0.30	41.82	0.50	74.62	0.03	21.94	96.59	
						W	0.30	∞	0.50				96.93	
2	9	3.16	4.55	0.36	50	SPE	−62.83	43.51	0.50	0.04	0.03	20.25	20.32	219.75
						SEQPE	0.23	43.51	0.50	69.95	0.03	20.25	90.23	
						W	0.22	∞	0.50				90.55	
3	9	3.16	4.55	0.64	50	SPE	−68.24	38.64	0.50	0.06	0.03	25.68	25.77	227.83
						SEQPE	0.13	38.64	0.50	92.63	0.03	25.68	118.35	
						W	0.13	∞	0.50				118.76	
4	9	3.16	4.55	0.45	50	SPE	−64.96	42.42	0.50	0.05	0.03	21.42	21.50	232.49
						SEQPE	0.18	42.42	0.50	78.17	0.03	21.41	99.61	
						W	0.18	∞	0.50				99.95	
5	9	3.16	4.55	0.55	50	SPE	−66.78	40.91	0.50	0.07	0.03	23.07	23.17	236.69
						SEQPE	0.15	40.91	0.50	85.89	0.03	23.07	108.99	
						W	0.15	∞	0.50				109.35	

Note: $x_2^{SPE} = 1$ and $x_2^{SEQPE} = 1$ omitted.

Proposition 8. Suppose $\frac{a-c}{b} > \frac{7}{2}$ and $\gamma' > \gamma$. If $x_1^{SPE}(\gamma') = x_1^{SPE}(\gamma)$, then $m(t_1^{SPE}(\gamma')) < m(t_1^{SPE}(\gamma))$.

Note that Proposition 8 slightly strengthens our maintained assumption from Section 2 that $\frac{a-c}{b} > 3$. Proposition 8 says that under SPE, conditional on the location of first entry, increasing the rate of market growth decreases the threshold for market size at the time of first entry.

Building on equation (30) and Proposition 8, we test whether SEQPE or SPE better fits the data by regressing market size at the time of first entry $m(t_1)$ on the market growth rate γ while controlling for the location of first entry x_1. This test faces two key challenges. First, as equation (30) makes clear, we have to measure x_1 and how x_1 maps into profits, as well as r and F. This measurement has to be consistent across the markets in our data. Second, the location of first entry is codetermined with market size at the time of first entry. Anything that we cannot measure potentially affects both x_1 on the right-hand side and $m(t_1)$ on the left, thereby creating an endogeneity problem. In our application, we argue that whatever we cannot measure and explicitly control for is either absorbed by fixed effects or can be addressed with instrumental variables.

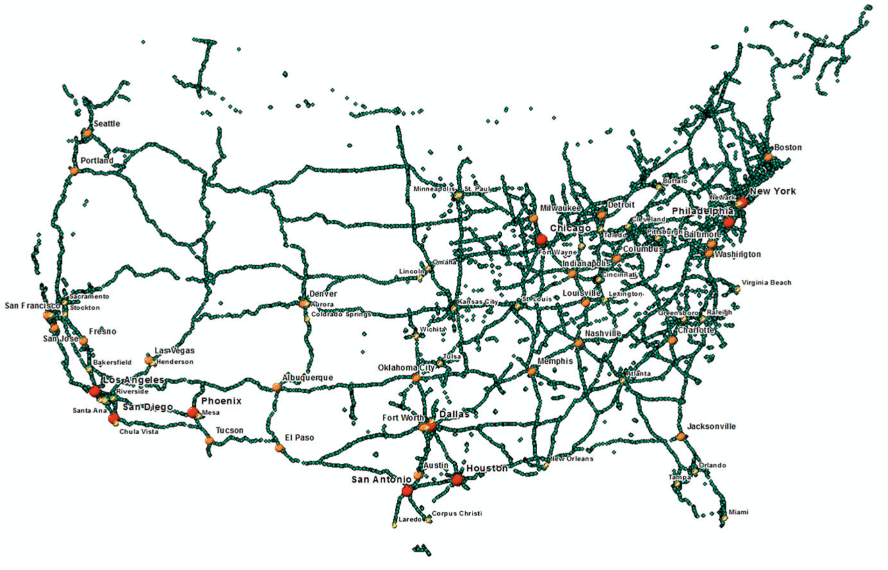
Although more than two firms may eventually enter a market in our empirical application, we expect the qualitative predictions of our model to remain unchanged. If a second entrant has to account for preemption by another firm under SPE, then this will accelerate entry by the second entrant, which in turn will accelerate entry by the first entrant.

Data and estimation. We study the timing and location of first entry by gas stations, restaurants, and hotels in geographically isolated markets around highway exits and intersections. We choose these industries because they vary in the fixed cost of entry and in the extent to which spatial differentiation matters for profitability.

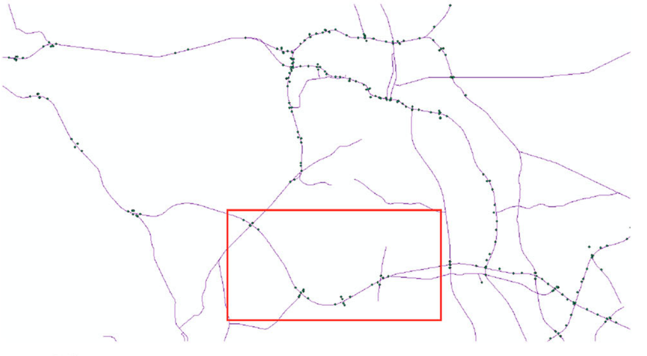
We use the ESRI Data and Maps 2013 data available through ArcGIS at www.esri.com to define markets. The data record the latitude and longitude of all highway exits and intersections in the United States. We visualize the data in panel (a) of Figure 3. To define a market, we group together any road crossings (i.e., highway exits or intersections) that are within 1500 m of each other to form a cluster. Each cluster is a market. Panels (b) and (c) illustrate this construction in a close-up of Durham, NC. In panel (c), we draw circles with radius 1500 m around all road crossings. If a road crossing lies in the circle drawn around another road crossing, then the two road crossings belong to the same market. In this example, there are four distinct markets, shown

FIGURE 3

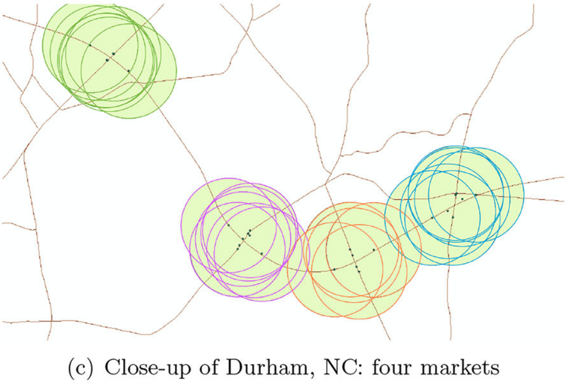
MARKET DEFINITION



(a) Highway exits and intersections



(b) Close-up of Durham, NC: exits and intersections



(c) Close-up of Durham, NC: four markets with color-coded circles. We define the center of a market as the average latitude and longitude of the road crossings in that market. To avoid urban and sprawling suburban areas without a clear market center, we drop clusters with width or height exceeding 2500 m. In urban areas, sequential highway exits and intersections tend to be narrowly spaced and thus tend to form large clusters. By contrast, the elongated clusters along interstate highways typically fall within the limit of 2500 m.

We obtain the entry dates and locations of gas stations, restaurants, and hotels in each market from Reference USA's Historical Business Database for 1997–2006. We use each year's database to avoid selection bias, which would be a concern if we used the entry dates and locations for only the current set of firms. We designate as a first entrant a firm that enters a market that had no firms of that type in the previous year. The sample contains 704 markets with a first entry by a gas station, 875 markets with a first entry by a restaurant, 527 markets with a first entry by a hotel, and 534 markets with a first entry by a three-star hotel.[12] Out of these, we have 664 markets with a second entry by a gas station, 774 markets with second entry by a restaurant, 494 markets with a second entry by a hotel, and 383 markets with a second entry by a three-star hotel.

For each firm, we calculate its Euclidian distance to the market center. In our model, the first entrant does not locate farther away from the market center than the second entrant. For those markets in our data with a second entrant, the distance between the market center and the first entrant is more than 10% larger than the distance between the market center and the second entrant in only 6%, 7%, 9%, and 14% of markets for gas stations, restaurants, hotels, and three-star hotels, respectively.[13]

Turning to market size and the rate of market growth, we use highway traffic data that we obtain directly from the Department of Transportation for 1993–2009, the last year the data were available.[14] Using m to index markets and t to index years, we measure market size S_{mt} by annual average daily traffic (AADT) and the rate of market growth γ_{mt} by the annualized growth of AADT between $t-2$ and $t+2$ (i.e., $\gamma_{mt} = \left(1 + \frac{S_{mt+2} - S_{mt-2}}{S_{mt}}\right)^{\frac{1}{4}} - 1$). We provide details on how we allocate AADT to markets in online Appendix E.

As our focus is on the relationship between market size at the time of first entry and the rate of market growth, a potential concern is that market size and the rate of market growth may be correlated more generally for a variety of reasons. For example, larger markets may experience slower growth due to capacity constraints and congestion of highways. Moreover, Proposition 8 assumes an exponential market size function with a constant rate of market growth, which implies that market size is unrelated to the rate of market growth. To investigate any underlying correlation between market size and the market growth rate, we use a balanced panel of markets in which we observe at least one gas station, restaurant, hotel, or three-star hotel for 1993–2006. The final sample includes 259 first entry events for gas stations, 277 for restaurants, 163 for hotels, and 156 for three-star hotels. In Figure 4, we plot the rate of market growth against the log of market size using a binscatter plot with 20 bins of equal size. Although there is a negative correlation between size and the growth rate for smaller markets, it disappears for larger markets. We therefore restrict the subsequent analyses to markets with AADT exceeding 8000 at the time of first entry.[15] With this restricted sample ($N = 67,890$), we regress the rate of market growth on the log of market size and state and year fixed effects, using two-way clustering on state and year. We do not find a statistically significant relationship between the market growth rate and market size.

We assume that the rate of market growth is not affected by the entry of the firms we study. That is, we assume that highway traffic does not increase simply because of a gas station, restau-

[12] When examining three-star hotels, we ignore hotels of other star levels that may have entered earlier.

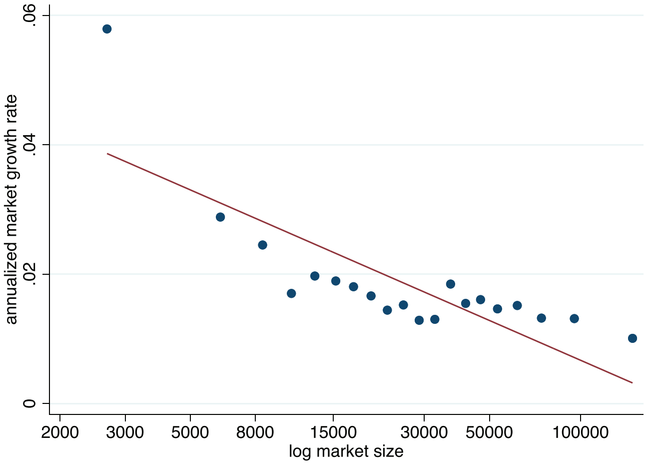
[13] We allow ourselves a 10% "buffer" in this exercise due to spatial constraints on entry locations.

[14] The Federal Highway Administration switched from raw file format in the earlier years to shapefile format in the later years. As a result, there are no data for 2010.

[15] Our results are robust to restricting the analysis further to markets with AADT in excess of 8000 for all years.

FIGURE 4

RATE OF MARKET GROWTH AND MARKET SIZE



rant, or hotel, which we consider reasonable. Even with this assumption, we must still account for factors such as entry costs that are likely to affect both the location of first entry and the market's size at that time. Beyond including dummies for the number of highways in the market, the type of the largest road,[16] and state and year fixed effects, we collect data on land prices in 2018 from the American Enterprise Institute and on vegetation in 2001 from the United States Geological Survey LANDFIRE Data Distribution Site. The latter data are typically used to assess fire risk, but we use them as shifters of entry costs, as it costs more to build a new establishment in a dense forest than in a field. Finally, we collect information about the number of other businesses in the market within 500 and 1000 m of the market center, which may reflect entry costs at these distances. We provide summary statistics and further details on how we assembled the data in online Appendix E.

In addition to these controls, we construct two instruments for the location of first entry to address the aforementioned endogeneity concern. Our instruments capture the geographic particularities of a market that may shift the location of first entry. First, we use the log of the diagonal of the bounding box of the market, determined by the maximum difference in the latitudes and longitudes of the road crossings in the market, to measure how geographically spread out the road crossings are. We expect more spread out road crossings to push firms farther out. We interact this instrument with dummies for Michigan, Ohio, Texas, North Carolina, Mississippi, and Georgia, because these states are most represented in the data. This allows the geographic size of the market (i.e., how spread out it is) to have a different sized effect on first entrant's location depending on the state. In the second IV regression, we also use the log distance between the market center and the eventual second entrant, which may also reflect geographic particularities of the market.

To test the relationship between market size at the time of first entry and the rate of market growth, we estimate the model

$$\log S_{mt} = f(\gamma_{mt}) + \alpha \log D_{mt} + X_m\beta + \zeta_{s(m)} + \xi_t + \epsilon_{mt}, \tag{31}$$

where S_{mt} is market size at the time of first entry, $f(\gamma_{mt})$ is a nonparametric function of the rate of market growth at the time of first entry, D_{mt} is the distance between the market center and the

[16] The types are interstate, other freeways and expressways and principal arterials, and other (minor arterial, major collector, minor collector, and local).

first entrant, and X_m is a vector of controls that includes an intercept, dummies for the number of highways, dummies for the type of the largest road, the number of other businesses within 500 and 1000 m, and variables for land prices and vegetation (we combine shrub and herb with forest). We also include state fixed effects $\zeta_{s(m)}$, where $s(m)$ is the state in which market m is located, and year fixed effects ξ_t.[17]

We expect $f(\gamma_{mt})$ to be a constant under SEQPE and a decreasing function under SPE. To construct the nonparametric function, we bin the rate of market growth in ranges of 3%. To facilitate estimating equation (31) alternatively by ordinary least squares (OLS) and instrumental variable (IV), we replace $f(\gamma_{mt})$ by a linear function of γ_{mt} for our main results in Tables 7–11. We exclude from the estimation those markets that have a negative growth at the time of first entry. Throughout, we proceed separately for gas stations, restaurants, hotels, and three-star hotels.

Results. We plot the estimated $f(\gamma_{mt})$ in Figure 5 for our four types of firms. For gas stations and three-star hotels, we see a declining relationship between market size at the time of first entry and the rate of market growth, consistent with SPE. However, the error bars are large, especially at higher rates of market growth where we have fewer observations. In contrast, we do not see a clear decline in market size at the time of first entry for restaurants, although there is a decline in point estimate moving from 6% to 9% market growth to >9% growth. For hotels, the effect is flat until moving from 6% to 9% market growth to >9% growth, when the point estimate actually increases.

In column (1) of Table 7, we show the OLS estimates for gas stations when replacing the flexible function of growth rate with a linear effect. We find a statistically significant negative coefficient (at 10% with a two-sided test) on γ_{mt}, consistent with SPE. The two IV regressions then follow in columns (2)–(5), with the first stages in columns (2) and (4) and the second stages in columns (3) and (5). Once again, we find a statistically significant negative impact of the rate of market growth (10% and 5%, respectively), with the coefficients not significantly different from the OLS estimates. As expected, in the first stage of the first IV regression, the diagonal of the bounding box of the market has a statistically significant positive impact on the distance between the market center and the first entrant, and this relationship is stronger and weaker for different states. Because we cluster standard errors by state, we report the Kleibergen–Paap first stage F statistic for the instruments, showing that we do not have weak instruments Kleibergen (2002). We also report the Hansen J statistic for the test of over-identifying restrictions, which we pass easily. In the first stage of the second IV regression, we find a statistically significant positive impact of the log distance between the market center and the second entrant on the log distance between the market center and the first entrant. This instrument is very strong, leading to a Kleibergen–Paap first stage F statistic of over 800.

As expected from Figure 5 results, we do not find a statistically significant relationship between market size at the time of first entry and the rate of market growth for restaurants in Table 8 and hotels in Table 9. This is notable for two reasons. First, using highway exits and intersections to define a market is likely most relevant for gas stations, whereas restaurants and hotels may compete on a wider geographic scale. Second, gas stations have less scope for product differentiation than restaurants and hotels, and thus spatial differentiation is likely more important.

To home in on spatial differentiation and mitigate the impact of other dimensions, we again repeat our analysis for three-star hotels. The OLS and IV estimates in Table 10 show a statistically significant negative impact of the rate of market growth (at 5% for all three specifications), again consistent with SPE.

As an additional check of whether the negative relationship that we find between market size at the time of first entry and the rate of market growth is not spurious, we conduct a placebo test, where we randomly assign the entry date for the first entrant to another year between 1998 and

[17] Our extensive set of controls reduces the useable sample size. For example, states with only one market are fully explained by the state fixed effect.

TABLE 7 Regression Results for Gas Stations

Variable	OLS Log Market Size	IV 1: 1st Stage: Log Distance	IV 1: 2nd Stage: Log Market Size	IV 2: 1st Stage: Log Distance	IV 2: 2nd Stage: Log Market Size
Rate of market growth	−1.6603*	−1.8313	−1.3356*	0.1492	−1.6646**
	(0.8865)	(1.2771)	(0.7758)	(0.2395)	(0.8390)
Log distance to center (m)	−0.0955**		0.0499		−0.0974**
	(0.0437)		(0.1143)		(0.0390)
Land price index	0.0453	0.2399	0.0104	0.0854	0.0458
	(0.2972)	(0.2761)	(0.3073)	(0.0941)	(0.2809)
Largest road interstate highway	0.7539**	0.7124	0.5697*	0.0967	0.7564**
	(0.3378)	(0.4598)	(0.3176)	(0.1052)	(0.3163)
Largest road freeway/principle arterial	0.3684**	−0.4512*	0.4465***	−0.0068	0.3674**
	(0.1725)	(0.2270)	(0.1514)	(0.0272)	(0.1662)
Three-road intersection	0.0630	0.2587**	0.0215	−0.0248	0.0636
	(0.0987)	(0.1120)	(0.0932)	(0.0173)	(0.0934)
Four-road intersection	0.2153	0.0374	0.1933	−0.0126	0.2156
	(0.1454)	(0.1350)	(0.1373)	(0.0406)	(0.1370)
Five-road intersection	0.3685**	0.3060**	0.2988*	0.0083	0.3695**
	(0.1585)	(0.1177)	(0.1664)	(0.0580)	(0.1504)
Log diagonal of bounding box (m)		0.3572*		−0.0347	
		(0.1891)		(0.0554)	
MI X log diagonal of bounding box (m)		−0.4749**		−0.1163**	
		(0.1972)		(0.0498)	
OH X log diagonal of bounding box (m)		0.8688***		0.0381	
		(0.2539)		(0.0763)	
TX X log diagonal of bounding box (m)		−0.3850*		−0.0112	
		(0.1972)		(0.0527)	
NC X log diagonal of bounding box (m)		−0.1556		0.0079	
		(0.1981)		(0.0578)	
MS X log diagonal of bounding box (m)		0.6081**		−0.0844	
		(0.2970)		(0.0698)	
GA X log diagonal of bounding box (m)		1.2789***		−0.0038	
		(0.2362)		(0.0599)	
Log distance of second entrant to center (m)				1.0114***	
				(0.0211)	
Number of firms within 500 m	0.0021	0.0010	0.0018	−0.0003	0.0021
	(0.0029)	(0.0030)	(0.0027)	(0.0006)	(0.0028)
Number of firms within 1000 m	0.0016	−0.0023	0.0020	0.0003	0.0015
	(0.0015)	(0.0016)	(0.0015)	(0.0004)	(0.0014)
Forest coverage 500 m circle (sq. km)	−0.9128	1.6131	−1.2056*	−0.5134	−0.9089
	(0.6151)	(1.0717)	(0.6918)	(0.3561)	(0.5812)
Forest coverage 1000 m circle (sq. km)	0.0516	−0.1946	0.0967	0.1713	0.0510
	(0.1890)	(0.3185)	(0.2002)	(0.1017)	(0.1793)
Field/grass coverage 500 m circle (sq. km)	−0.7284	−0.5975	−0.6763	−0.1480	−0.7291
	(0.5904)	(1.0525)	(0.6224)	(0.2212)	(0.5576)
Field/grass coverage 1000 m circle (sq. km)	−0.1001	0.1163	−0.1093	0.0724	−0.0999
	(0.2114)	(0.2831)	(0.2148)	(0.0778)	(0.2000)
Water coverage 500 m circle (sq. km)	−0.3888	0.5812	−0.5444	0.3373	−0.3867
	(1.4231)	(1.9522)	(1.3645)	(0.6550)	(1.3530)
Water coverage 1000 m circle (sq. km)	0.2305	−0.0371	0.2508	−0.0617	0.2302
	(0.4461)	(0.7047)	(0.4293)	(0.1537)	(0.4236)
R^2	0.240	0.234	0.206	0.967	0.240
N	272	272	268	272	268
Kleibergen–Paap F statistic		73.53		1263.99	
Hansen J statistic		4.781		5.65	

Note: $p < .10$ (*), $p < .05$ (**), $p < .01$ (***). Standard errors clustered by state.

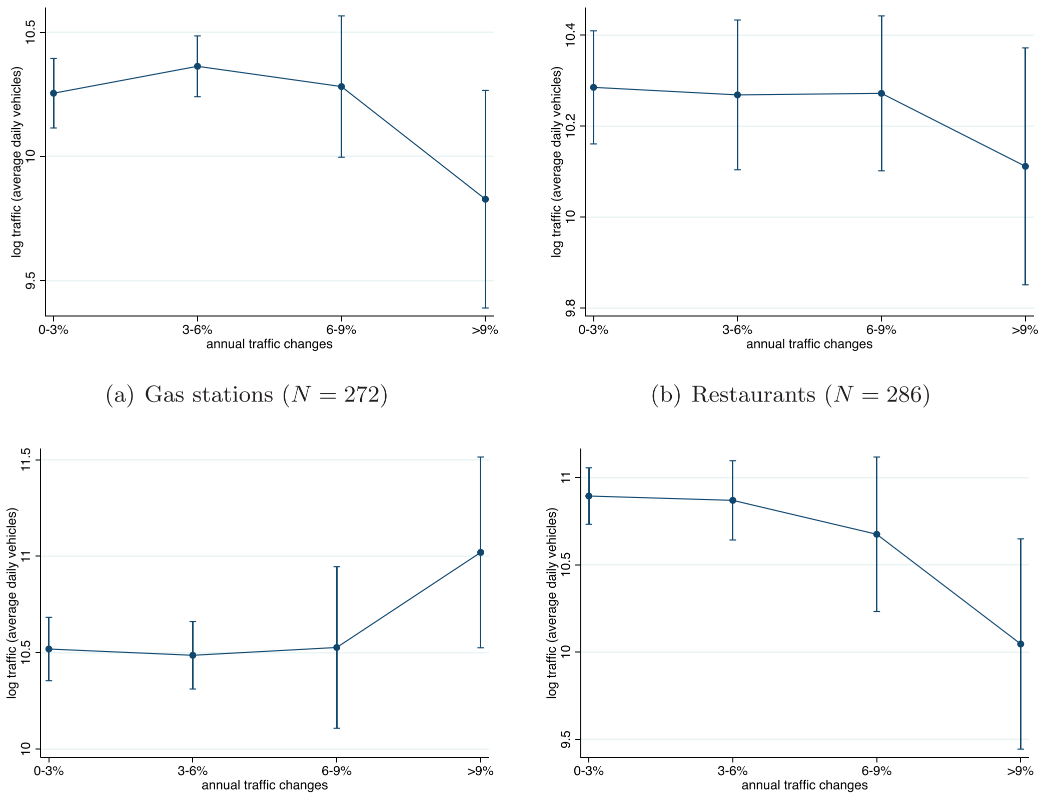
TABLE 11 Placebo Results for Gas Stations

Variable	OLS Log Market Size	IV 1: 1st Stage: Log Distance	IV 1: 2nd Stage: Log Market Size	IV 2: 1st Stage: Log Distance	IV 2: 2nd Stage: Log Market Size
Rate of market growth	−1.0922	2.5708**	−0.7747	0.0661	−1.0681
	(1.0928)	(1.1749)	(1.1386)	(0.2461)	(1.0344)
Log distance to center (m)	−0.0197		−0.1538		−0.0299
	(0.0446)		(0.1795)		(0.0411)
Land price index	0.0623	−0.1152	0.0421	−0.0275	0.0608
	(0.1109)	(0.1371)	(0.1048)	(0.0212)	(0.1047)
Largest road interstate highway	0.5830**	0.4469	0.7407***	0.1005	0.5950**
	(0.2538)	(0.3095)	(0.2504)	(0.0775)	(0.2382)
Largest road freeway/principle arterial	0.3770**	−0.1139	0.3621**	−0.0363	0.3759***
	(0.1522)	(0.1776)	(0.1564)	(0.0365)	(0.1441)
Three-road intersection	0.1153	0.2178	0.1509	−0.0176	0.1180
	(0.1088)	(0.1342)	(0.1233)	(0.0164)	(0.1019)
Four-road intersection	0.2055**	−0.0236	0.2130**	−0.0435	0.2061**
	(0.0982)	(0.1783)	(0.0937)	(0.0280)	(0.0925)
Five-road intersection	0.2992**	0.4604***	0.3783**	−0.0604	0.3052**
	(0.1470)	(0.1571)	(0.1900)	(0.0409)	(0.1389)
Log diagonal of bounding box (m)		0.4856***		−0.0599**	
		(0.1464)		(0.0270)	
MI X log diagonal of bounding box (m)		−0.9095***		−0.0708**	
		(0.1777)		(0.0324)	
OH X log diagonal of bounding box (m)		0.4411*		0.0955*	
		(0.2242)		(0.0548)	
TX X log diagonal of bounding box (m)		−0.2355*		−0.0298	
		(0.1317)		(0.0286)	
NC X log diagonal of bounding box (m)		0.3345*		0.0859***	
		(0.1953)		(0.0288)	
MS X log diagonal of bounding box (m)		0.2133		−0.0258	
		(0.3791)		(0.0818)	
GA X log diagonal of bounding box (m)		−0.3786		0.1557**	
		(0.2869)		(0.0612)	
Log distance of second entrant to center (m)				1.0308***	
				(0.0140)	
Number of firms within 500 m	0.0025	−0.0042	0.0022	−0.0010	0.0025
	(0.0027)	(0.0050)	(0.0027)	(0.0007)	(0.0025)
Number of firms within 1000 m	0.0032**	−0.0008	0.0030***	0.0004	0.0031***
	(0.0012)	(0.0019)	(0.0011)	(0.0002)	(0.0011)
Forest coverage 500 m circle (sq. km)	−0.4546	−0.1981	−0.4912	−0.2718	−0.4574
	(0.6940)	(0.9196)	(0.7210)	(0.1806)	(0.6599)
Forest coverage 1000 m circle (sq. km)	0.0509	0.0195	0.0607	0.0784	0.0516
	(0.2401)	(0.2247)	(0.2390)	(0.0644)	(0.2275)
Field/grass coverage 500 m circle (sq. km)	−0.6436	0.5433	−0.6050	−0.2055	−0.6407
	(0.9205)	(0.9000)	(0.9390)	(0.1903)	(0.8736)
Field/grass coverage 1000 m circle (sq. km)	0.0463	−0.1223	0.0356	0.0368	0.0455
	(0.2875)	(0.2475)	(0.2891)	(0.0630)	(0.2728)
Water coverage 500 m circle (sq. km)	−0.4513	2.9466*	−0.0333	0.5285	−0.4196
	(1.2044)	(1.5148)	(1.2653)	(0.5778)	(1.1498)
Water coverage 1000 m circle (sq. km)	0.2991	−0.9328*	0.1924	−0.1935	0.2910
	(0.3169)	(0.5326)	(0.3390)	(0.1487)	(0.3039)
R^2	0.177	0.227	0.150	0.972	0.177
N	279	279	278	279	278
Kleibergen–Paap F statistic		42.40		1627.38	
Hansen J statistic		4.76		5.54	

Note: $p < .10$ (*), $p < .05$ (**), $p < .01$ (***). Standard errors clustered by state.

FIGURE 5

MARKET SIZE THRESHOLD AND RATE OF MARKET GROWTH



(a) Gas stations ($N = 272$)
(b) Restaurants ($N = 286$)
(c) Hotels ($N = 196$)
(d) Three-star hotels ($N = 174$)

TABLE 8 Regression Results for Restaurants

Variable	OLS Log Market Size	IV 1		IV 2	
		1st Stage: Log Distance	2nd Stage: Log Market Size	1st Stage: Log Distance	2nd Stage: Log Market Size
Rate of market growth	−1.0371	0.8568	−1.1947	−0.0800	−1.0389
	(1.1491)	(1.8118)	(1.1623)	(0.2842)	(1.0881)
Log distance to center (m)	−0.0340		0.0621		−0.0329
	(0.0458)		(0.1975)		(0.0444)
Land price index	−0.1003	−0.3250	−0.0684	0.0264	−0.1000
	(0.2375)	(0.3199)	(0.2302)	(0.0387)	(0.2260)
Largest road interstate highway	0.9302***	0.6281**	0.8141*	0.0530	0.9289***
	(0.3216)	(0.2858)	(0.4315)	(0.0404)	(0.3065)
Largest road freeway/principle arterial	0.5511***	−0.0709	0.5525***	−0.0153	0.5511***
	(0.1154)	(0.1815)	(0.1081)	(0.0212)	(0.1096)
Three-road intersection	0.1684*	−0.1442	0.1757*	0.0079	0.1685*
	(0.0965)	(0.1336)	(0.1018)	(0.0218)	(0.0916)
Four-road intersection	0.2850***	0.1235	0.2634***	0.0122	0.2847***
	(0.1006)	(0.1956)	(0.0911)	(0.0305)	(0.0956)
Five-road intersection	0.3064**	−0.0163	0.2927**	−0.0184	0.3062***
	(0.1222)	(0.1967)	(0.1210)	(0.0277)	(0.1157)
Log diagonal of bounding box (m)		0.2568*		−0.0419	
		(0.1270)		(0.0300)	
MI X log diagonal of bounding box (m)		0.1266		−0.0763*	
		(0.2125)		(0.0399)	
OH X log diagonal of bounding box (m)		0.0025		−0.0392	
		(0.2548)		(0.0507)	
TX X log diagonal of bounding box (m)		0.1253		0.0215	
		(0.1346)		(0.0191)	
NC X log diagonal of bounding box (m)		0.3621**		−0.0259	
		(0.1538)		(0.0356)	
MS X log diagonal of bounding box (m)		2.0985***		0.1302	
		(0.4324)		(0.0881)	
GA X log diagonal of bounding box (m)		0.0530		−0.0625*	
		(0.1933)		(0.0311)	
Log distance of second entrant to center (m)				1.0007***	
				(0.0152)	
Number of firms within 500 m	−0.0061	−0.0558***	0.0001	−0.0022	−0.0060
	(0.0145)	(0.0158)	(0.0192)	(0.0026)	(0.0141)
Number of firms within 1000 m	0.0056*	0.0004	0.0055*	0.0005	0.0056*
	(0.0032)	(0.0031)	(0.0031)	(0.0008)	(0.0031)
Forest coverage 500 m circle (sq. km)	−0.0036	0.9972	−0.1194	0.2059	−0.0049
	(0.6884)	(0.8599)	(0.7752)	(0.1800)	(0.6561)
Forest coverage 1000 m circle (sq. km)	−0.1141	−0.2372	−0.0872	−0.0290	−0.1138
	(0.1612)	(0.1963)	(0.1730)	(0.0592)	(0.1540)
Field/grass coverage 500 m circle (sq. km)	−0.4654	3.1862**	−0.7397	0.1585	−0.4686
	(0.6112)	(1.2555)	(0.8535)	(0.2329)	(0.5823)
Field/grass coverage 1000 m circle (sq. km)	−0.0444	−0.8408**	0.0262	−0.0261	−0.0436
	(0.1806)	(0.3237)	(0.2357)	(0.0601)	(0.1734)
Water coverage 500 m circle (sq. km)	−3.3385**	−1.7108	−3.1573**	0.2815	−3.3364**
	(1.5710)	(2.0903)	(1.5294)	(0.5546)	(1.4935)
Water coverage 1000 m circle (sq. km)	0.5949*	−0.2055	0.6143*	−0.0512	0.5951*
	(0.3442)	(0.4178)	(0.3197)	(0.1299)	(0.3272)
R^2	0.180	0.268	0.168	0.978	0.180
N	286	286	281	286	281
Kleibergen–Paap F statistic		29.49		1922.07	
Hansen J statistic		5.51		6.25	

Note: $p < .10$ (*), $p < .05$ (**), $p < .01$ (***). Standard errors clustered by state.

TABLE 9 Regression Results for Hotels

		IV 1		IV 2	
Variable	OLS Log Market Size	1st Stage: Log Distance	2nd Stage: Log Market Size	1st Stage: Log Distance	2nd Stage: Log Market Size
Rate of market growth	−2.2372	1.8441	−2.3031	−0.0291	−2.2280
	(1.5468)	(1.3285)	(1.6976)	(0.2154)	(1.4466)
Log distance to center (m)	0.0074		0.0393		0.0029
	(0.0609)		(0.3297)		(0.0585)
Land price index	0.1558	−0.6666***	0.1768	0.0436	0.1528
	(0.1796)	(0.1911)	(0.2584)	(0.0308)	(0.1665)
Largest road interstate highway	1.0368***	0.4164	1.0077***	0.1768***	1.0409***
	(0.2306)	(0.3461)	(0.3685)	(0.0497)	(0.2153)
Largest road freeway/principle arterial	0.5175**	−0.1277	0.5230**	0.0464	0.5168**
	(0.2317)	(0.2562)	(0.2235)	(0.0351)	(0.2170)
Three-road intersection	0.3450**	−0.0200	0.3438***	0.0208	0.3452***
	(0.1278)	(0.1440)	(0.1208)	(0.0256)	(0.1191)
Four-road intersection	0.3840**	0.4291***	0.3696	0.0270	0.3860**
	(0.1891)	(0.1311)	(0.2524)	(0.0273)	(0.1759)
Five-road intersection	0.6594***	0.0440	0.6517***	−0.0094	0.6604***
	(0.1550)	(0.1620)	(0.1764)	(0.0369)	(0.1451)
Log diagonal of bounding box (m)		0.3396**		−0.0309	
		(0.1426)		(0.0200)	
MI X log diagonal of bounding box (m)		−0.1871		0.0161	
		(0.1713)		(0.0252)	
OH X log diagonal of bounding box (m)		0.2824		0.0462	
		(0.2102)		(0.0443)	
TX X log diagonal of bounding box (m)		−0.0973		0.0503***	
		(0.1359)		(0.0155)	
NC X log diagonal of bounding box (m)		0.2269		0.0213	
		(0.1397)		(0.0311)	
MS X log diagonal of bounding box (m)		0.7116***		0.0297	
		(0.1701)		(0.0482)	
GA X log diagonal of bounding box (m)		0.2984		0.0908***	
		(0.2442)		(0.0298)	
Log distance of second entrant to center (m)				0.9726***	
				(0.0168)	
Number of firms within 500 m	−0.0013	0.0013	−0.0013	0.0000	−0.0013
	(0.0029)	(0.0021)	(0.0026)	(0.0003)	(0.0027)
Number of firms within 1000 m	0.0004	−0.0004	0.0004	0.0000	0.0004
	(0.0016)	(0.0008)	(0.0014)	(0.0001)	(0.0014)
Forest coverage 500 m circle (sq. km)	0.0651	2.5703***	−0.0345	0.1007	0.0792
	(0.8269)	(0.8542)	(1.4368)	(0.1450)	(0.7664)
Forest coverage 1000 m circle (sq. km)	−0.1543	−0.4567**	−0.1356	−0.0150	−0.1570
	(0.1738)	(0.2078)	(0.2453)	(0.0400)	(0.1609)
Field/grass coverage 500 m circle (sq. km)	1.4208*	0.0364	1.4066**	−0.2373	1.4227**
	(0.7551)	(1.2515)	(0.7094)	(0.2729)	(0.7044)
Field/grass coverage 1000 m circle (sq. km)	−0.3877	−0.0361	−0.3841*	0.0824	−0.3882*
	(0.2520)	(0.3267)	(0.2293)	(0.0601)	(0.2347)
Water coverage 500 m circle (sq. km)	0.6240	−0.1184	0.6038	0.1287	0.6269
	(1.8309)	(2.5191)	(1.7525)	(0.6004)	(1.7059)
Water coverage 1000 m circle (sq. km)	−0.1016	0.1842	−0.1069	0.0992	−0.1008
	(0.3919)	(0.4435)	(0.3904)	(0.1117)	(0.3657)
R^2	0.241	0.403	0.240	0.984	0.241
N	196	196	193	196	193
Kleibergen–Paap F statistic		19.27		2272.74	
Hansen J statistic		6.75		9.03	

Note: $p < .10$ (*), $p < .05$ (**), $p < .01$ (***). Standard errors clustered by state.

TABLE 10 Regression Results for Three-star Hotels

		IV 1		IV 2	
Variable	OLS Log Market Size	1st Stage: Log Distance	2nd Stage: Log Market Size	1st Stage: Log Distance	2nd Stage: Log Market Size
Rate of market growth	−1.7374**	−1.3857**	−1.8771**	−0.0371	−1.8299**
	(0.7858)	(0.5758)	(0.7677)	(0.2831)	(0.7162)
Log distance to center (m)	−0.1208		−0.2287		−0.1923*
	(0.0934)		(0.2229)		(0.0998)
Land price index	0.5909*	−0.2779	0.5623*	0.3756*	0.5719*
	(0.3304)	(0.4562)	(0.3252)	(0.2096)	(0.3082)
Largest road interstate highway	0.4549	−0.2047	0.5492	0.1431	0.5174
	(0.5095)	(0.5067)	(0.5704)	(0.1391)	(0.4784)
Largest road freeway/principle arterial	0.2947	−0.5132	0.2476	0.0049	0.2635
	(0.5134)	(0.3683)	(0.4219)	(0.1260)	(0.4632)
Three-road intersection	0.0310	0.0172	0.0406	0.0019	0.0374
	(0.1806)	(0.1163)	(0.1709)	(0.0627)	(0.1678)
Four-road intersection	0.0614	0.0989	0.0903	−0.0378	0.0805
	(0.1950)	(0.2141)	(0.1624)	(0.0643)	(0.1760)
Five-road intersection	0.1514	0.1710	0.2016	0.0781	0.1847
	(0.2476)	(0.1570)	(0.2446)	(0.0688)	(0.2273)
Log diagonal of bounding box (m)		0.5476***		−0.0688	
		(0.1293)		(0.0660)	
MI X log diagonal of bounding box (m)		0.0816		0.1419***	
		(0.1516)		(0.0393)	
OH X log diagonal of bounding box (m)		−0.1412		−0.0570	
		(0.1099)		(0.0776)	
TX X log diagonal of bounding box (m)		−0.0753		0.0435	
		(0.1340)		(0.0459)	
NC X log diagonal of bounding box (m)		0.2388		0.0165	
		(0.1789)		(0.0615)	
MS X log diagonal of bounding box (m)		0.9509**		0.2212*	
		(0.3599)		(0.1279)	
GA X log diagonal of bounding box (m)		−0.0713		0.1081***	
		(0.0910)		(0.0378)	
Log distance of second entrant to center (m)				0.9243***	
				(0.0665)	
Number of firms within 500 m	−0.0018	−0.0021	−0.0020	−0.0005	−0.0019
	(0.0017)	(0.0032)	(0.0018)	(0.0007)	(0.0016)
Number of firms within 1000 m	0.0006	0.0001	0.0006	0.0001	0.0006
	(0.0004)	(0.0008)	(0.0004)	(0.0002)	(0.0004)
Forest coverage 500 m circle (sq. km)	1.5246	−0.8346	1.4507	0.1327	1.4756
	(1.0800)	(1.5022)	(0.9896)	(0.7187)	(0.9845)
Forest coverage 1000 m circle (sq. km)	−0.4308	0.3594	−0.4080	−0.0068	−0.4157
	(0.3207)	(0.3647)	(0.2900)	(0.1467)	(0.2936)
Field/grass coverage 500 m circle (sq. km)	3.5179*	−0.1188	3.5774**	−0.6064	3.5573**
	(1.8696)	(1.5523)	(1.7000)	(0.4293)	(1.7118)
Field/grass coverage 1000 m circle (sq. km)	−1.1475**	−0.0687	−1.1731***	0.0999	−1.1645***
	(0.4580)	(0.3692)	(0.4150)	(0.1173)	(0.4152)
Water coverage 500 m circle (sq. km)	−2.0660	−0.2568	−2.1651	−1.3025	−2.1317
	(1.5716)	(1.9443)	(1.4869)	(1.4209)	(1.4351)
Water coverage 1000 m circle (sq. km)	−0.0997	−0.2357	−0.1176	0.4347*	−0.1116
	(0.4878)	(0.3369)	(0.4312)	(0.2484)	(0.4472)
R^2	0.202	0.527	0.192	0.919	0.197
N	174	174	170	174	170
Kleibergen–Paap F statistic		7.84		110.13	
Hansen J statistic		7.39		7.413	

Note: $p < .10$ (*), $p < .05$ (**), $p < .01$ (***). Standard errors clustered by state.

2006. In the OLS and IV estimates for gas stations in Table 11, there is no longer a statistically significant impact of the rate of market growth. The same is true for the restaurants, hotels, and three-star hotels (not reported due to space constraints).

Our results for market size thresholds resemble those in Bresnahan and Reiss (1990), who consider a static entry game, first with simultaneous and then with sequential moves. They find that the breakeven market size of monopoly and duopoly markets for automobile dealers decreases with positive market growth (as measured by population growth) and increases with negative market growth. This further supports our conclusion that SPE can be a more natural equilibrium concept than SEQPE.

8. Conclusions

■ In this article, we contribute to a growing literature in empirical IO by developing and analyzing a dynamic model of entry around a model of spatial competition. In our model, firms decide when and where to enter a growing market. In a SEQPE, firms pre-commit to a time and location of entry, with preemption ruled out by assumption. By contrast, in a SPE, the threat of preemption is operative until rents equalize. In analyzing how this affects the timing and location of entry, we show that the threat of preemption leads to premature entry, less extreme locations, and the dissipation of rents.

In our empirical application, we apply this model to study the timing and location of first entry by gas stations, restaurants, and hotels in geographically isolated markets around highway exits and intersections. We show that the nature of the game being played generates testable predictions for the relationship between market size at the time of first entry and the market growth rate. We find that gas stations are consistent with SPE but restaurants and hotels are not. This finding may be explained by institutional details, in particular the fact that restaurants and hotels combine spatial differentiation with other dimensions of product differentiation. Indeed, when we isolate spatial differentiation by examining three-star hotels, the data are again consistent with SPE. Taken together, we conclude that SPE can be a more natural equilibrium concept than SEQPE for understanding the evolution of growing markets.

References

Anderson, S., Goeree, J., and Ramer, R. "Location, Location, Location." *Journal of Economic Theory*, Vol. 77 (1997), pp. 102–127.

Arcidiacono, P., Bayer, P., Blevins, J., and Ellickson, P. "Estimation of Dynamic Discrete Choice Models in Continuous Time with an Application to Retail Competition." *Review of Economic Studies*, Vol. 83 (2016), pp. 889–931.

Argenziano, R. and Schmidt-Dengler, P. "Competition, Timing of Entry and Welfare in a Preemption Game." *Economics Letters*, Vol. 120 (2013), pp. 509–512.

Bresnahan, T. and Reiss, P. "Entry in Monopoly Markets." *Review of Economic Studies*, Vol. 57 (1990), pp. 531–553.

Butters, A. and Hubbard, T. "Industry Structure, Segmentation and Quality Competition in the U.S. Hotel Industry." *Journal of Industrial Economics*, Vol. 71 (2023), 491–537.

Collard-Wexler, A. "Demand Fluctuations in the Ready-Mix Concrete Industry." *Econometrica*, Vol. 81 (2013), pp. 1003–1037.

d'Aspremont, C., Gabszewicz, J., and Thisse, J. "On Hotelling's "Stability in Competition"." *Econometrica*, Vol. 47 (1979), pp. 1145–1150.

Economides, N. "The Principle of Minimum Differentiation Revisited." *European Economic Review*, Vol. 24 (1984), pp. 345–368.

Economides, N. "Minimal and Maximal Product Differentiation in Hotelling's Duopoly." *Economics Letters*, Vol. 21 (1986), pp. 67–71.

Ericson, R. and Pakes, A. "Markov-Perfect Industry Dynamics: A Framework for Empirical Work." *Review of Economic Studies*, Vol. 62 (1995), pp. 53–82.

Fang, L. and Yang, N. "Aggressive Growth in Retail: A Trade-off Between Deterrence and Survival?" Working Paper, University of British Columbia, Vancouver, 2019.

Fang, L. and Yang, N. "Measuring Preemptive Entry in Dynamic Oligopoly Games." Working Paper, University of British Columbia, Vancouver, 2022.

Fudenberg, D. and Tirole, J. "Preemption and Rent Equalization in the Adoption of New Technology." *Review of Economic Studies*, Vol. 52 (1985), pp. 383–401.

Gilbert, R. and Harris, R. "Competition with Lumpy Investment." *Rand Journal of Economics*, Vol. 15 (1984), pp. 197–212.

Hotelling, H. "Stability in Competition." *Economic Journal*, Vol. 39 (1929), pp. 41–58.

Igami, M. and Yang, N. "Unobserved Heterogeneity in Dynamic Games: Cannibalization and Preemptive Entry of Hamburger Chains in Canada." *Quantitative Economics*, Vol. 7 (2016), pp. 483–521.

Kleibergen, F. "Pivotal Statistics for Testing Structural Parameters in Instrumental Variables Regression." *Econometrica*, Vol. 70 (2002), pp. 1781–1803.

Lambertini, L. "Equilibrium Locations in a Spatial Model with Sequential Entry in Real Time." *Regional Science and Urban Economics*, Vol. 32 (2002), pp. 47–58.

Mazzeo, M. "Product Choice and Oligopoly Market Structure." *Rand Journal of Economics*, Vol. 33 (2002), pp. 221–242.

Neven, D. "Two Stage (perfect) Equilibrium in Hotelling's Model." *Journal of Industrial Economics*, Vol. 33 (1985), pp. 317–325.

Neven, D. "Endogenous Sequential Entry in a Spatial Model." *International Journal of Industrial Organization*, Vol. 5 (1987), pp. 419–434.

Reinganum, J. "Market Structure and the Diffusion of New Technology." *Bell Journal of Economics*, Vol. 12 (1981a), pp. 618–624.

Reinganum, J. "On the Diffusion of New Technology: A Game Theoretic Approach." *Review of Economic Studies*, Vol. 48 (1981b), pp. 395–405.

Riordan, M. "Regulation and Preemptive Technology Adoption." *Rand Journal of Economics*, Vol. 23 (1992), pp. 334–349.

Schmidt-Dengler, P. "The Timing of New Technology Adoption: The Case of MRI." Working Paper, University of Mannheim, Mannheim, 2006.

Shen, Q. and Villas-Boas, M. "Strategic Entry Before Demand Takes Off." *Management Science*, Vol. 56 (2010), pp. 1259–1271.

Simon, L. and Stinchcombe, M. "Extensive Form Games in Continuous Time: Pure Strategies." *Econometrica*, Vol. 57 (1989), pp. 1171–1214.

Smirnov, V. and Wait, A. "Innovation in a Generalized Timing Game." *International Journal of Industrial Organization*, Vol. 42 (2015), pp. 23–33.

Sweeting, A. "Dynamic Product Positioning in Differentiated Product Markets: The Effect of Fees for Musical Performance Rights on the Commercial Radio Industry." *Econometrica*, Vol. 81 (2013), pp. 1763–1803.

Tabuchi, T. and Thisse, J. "Asymmetric Equilibria in Spatial Competition." *International Journal of Industrial Organization*, Vol. 13 (1995), pp. 213–227.

Zheng, F. "Spatial Competition and Preemptive Entry in the Discount Retail Industry." Working Paper no. 16–37, Columbia University, New York, 2016.

Zhou, D. and Vertinsky, I. "Strategic Location Decisions in a Growing Market." *Regional Science and Urban Economics*, Vol. 31 (2001), pp. 523–533.

Supporting information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Figure 1: 2001 vegetation data. Blue = water, purple/red/black = developed, yellow = field and grass, green = forest, brown = shrub and herb
Table 1: 2001 vegetation data summary statistics.