

An Equilibrium Model of Investment Under Uncertainty

Robert Novy-Marx

University of Chicago

We analyze the optimal investment decisions of heterogeneous firms in a competitive, uncertain environment, characterizing firms' investment strategies explicitly and deriving closed-form solutions for firm value. Real option premia remain significant, and are even unmitigated relative to the standard partial-equilibrium model when both are calibrated to observables. Firms consequently delay investment, choosing not to undertake some positive NPV projects. We compare competitive behavior to that of a strategic monopolist, and quantify the welfare loss associated with monopoly. Finally, the model predicts business cycle dependence on firm returns, with returns negatively skewed during industry expansions but positively skewed in industry recessions. (*JEL* G12, D5, E22, R14)

Using real options theory to account for the value of future development is now standard in finance among both academics and practitioners. It provides a heuristic, intuitively appealing explanation of the real world's observed deviations from neoclassical Tobin's Q theory of investment. The most notable of these deviations are market values that exceed book values, often greatly, and investment thresholds that may significantly exceed zero net present value.

The theory's conclusion that delaying investment can result in excess profits seems to be at odds with what we know about competition. Competition drives down profit opportunities, a fact known long before Cournot modeled the effect in 1838.

This article presents an equilibrium model integrating the two disparate branches of the economics literature—real options theory and the theory of competition. Recent articles, most notably Grenadier (2002), have used the Cournot intuition to argue that competition erodes real option values and reduces investment delays.¹

I would like to thank Bob Anderson, Jonathan Berk, Jason Chen, Madhur Duggar, Janice Eberly, Steve Evans, Rob Gertner, Milena Novy-Marx, John Quigley, Mark Rubinstein, Jacob Sagi and Nancy Wallace for discussions and comments. All errors are mine alone. Address correspondence to Robert Novy-Marx, Graduate School of Business, University of Chicago, 5807 S Woodlawn Ave, Chicago IL, 60637, or e-mail: rnm@ChicagoGSB.edu.

¹ Several other articles have also considered the effects of competitive interactions on real options. Smets (1991), Grenadier (1996), Garlappi (2004), and Lambrecht and Perraudin (2003) have considered duopolistic settings and Leahy (1993) perfect competition; Spatt and Sterbenz (1985) and Williams (1993) have considered the intermediate cases.

These results are difficult to reconcile, however, with important sectors of the economy. Option premia are significantly positive and firms delay investment in some highly competitive industries. Titman (1985) illustrates this with a simple example: empty lots in city centers. Real estate is a highly competitive industry with many players, yet owners sometimes choose not to build on property that could certainly be developed profitably. Instead, as Titman argues, the value of a lot derives from the option to develop, and owners sometimes delay building believing that the lot can be developed more profitably later.

The analysis presented in this article shows that in a competitive industry firms can actually appear to deviate more from neoclassical behavior than the standard real options analysis predicts. In particular, firms may invest only at significantly positive option premia, and delay irreversible investment longer than predicted by a standard, partial-equilibrium model calibrated to the observed price data.

Grenadier's results—that competition erodes option values and pushes firms back to the zero net present value (NPV) investment rule—and similar results found in Leahy (1993) and Kogan (2001) are a consequence of the type of industry they model, one in which the production technology is linear and incremental.² In these articles firms may add capacity in arbitrarily small increments without suffering any adjustment costs. With this sort of production technology, real option premia are nothing more than rents to monopoly (or oligopoly), and it is hardly surprising then that sufficient competition destroys the option premium.

The prediction that competition drives option premia to zero is not, however, generally true. Oligopoly rents are not the only possible source of excess profit opportunities; rents may derive directly from the production technology. Option values and investment behavior in the real estate market example, and in many other industries, differ from Grenadier's predictions because undertaking investment often entails opportunity costs, and because firms vary in scope and size. We show in this article that in industries in which opportunity costs and heterogeneity are important real option values are significant, investment decisions are delayed, and investment is lumpy and intermittent. Moreover, this is true even when competition drives oligopoly rents to zero, because significant rents still accrue to the production technology.

Since Hotelling's (1929) seminal paper, it has been understood that demand side heterogeneity can reduce competition. The whole concept of horizontal product differentiation is predicated on the idea that variations in tastes allow firms to segment the market, compete less, and extract more consumer surplus. Heterogeneity can also provide a natural ordering to

² Similar results that also depend on the choice of an irreversible linear incremental production technology may also be found in Williams (1993) and Caballero and Pindyck (1996).

agents' actions. We do not all buy new computers, or cars, at the same time, at least in part because those of us with older, obsolete models are more likely near-term buyers than those with newer, contemporary models.

This article shows that supply side heterogeneity can reduce competition as well. When heterogeneity extends to costs or profitability, it is a wedge that breaks the idea of "perfect competition," even when firms are perfectly competitive. In the real estate example considered previously, a large number of firms compete vigorously, yet heterogeneity prevents them from all competing directly over any investment opportunity. When the owner of the empty lot considers putting up a 30-story office tower she is not competing with the owner of the 15-story apartment complex next door. The opportunity costs to the owner of the 15-story apartment complex, which include walking away from the existing building (i.e., tearing the building down and losing the revenue generated by the lot's existing improvements), effectively preclude her from competing with the owner of the empty lot. The owner of the 15-story apartment complex however still possesses a valuable option to compete over investment opportunities in the future. If the city grows, there may be demand in the future for a new 60-story office tower. If at that time all the empty lots are developed, the owner of the 15-story apartment complex may have low enough opportunity costs to compete.

While the fact that supply side heterogeneity reduces competition is not surprising, the heterogeneity itself is. Supply side heterogeneity reduces competition, but itself arises endogenously as a consequence of the fact that heterogeneous firms compete less. Firms naturally make investment decisions that differentiate them from others in a manner that reduces the amount of competition they face. The natural, stable cross-sectional distribution of firms is the one that minimizes intertemporal competition between firms. The degree of cross-sectional dispersion depends on the economic primitives. Greater cross-sectional variation in firms is associated with (1) high uncertainty regarding future demand, (2) a high average demand growth, (3) low cost-to-scale (i.e., close to linear) for developing capacity, and (4) a low discount rate.

We develop a model in which firms incur opportunity as well as direct investment costs to altering capacity, and face aggregate uncertainty regarding demand for their output, the price of which is determined endogenously and is a function of firms' investment decisions. We find that even with an infinite number of competitive heterogeneous firms, option premia are significant and firms optimally delay irreversible investment, choosing not to undertake some positive NPV projects.

The Cournot intuition that competition should drive firms to invest earlier is compelling, so our result that competition appears to lead firms to delay investment longer is somewhat surprising. Note, however, that we are not claiming that competitive agents delay longer than

a strategic monopolist facing the same demand side of the market, but that competitive firms delay investment longer than the standard, partial-equilibrium model predicts when calibrated to observed prices.

This result derives from the endogenously determined equilibrium price of firms' output. We show that the price of firms' output is negatively skewed because aggregate industry capacity responds asymmetrically to changing demand. Firms can add capacity quickly in response to rising demand, but cannot adjust capacity as quickly to falling demand owing to investment irreversibility. As a result, increasing supply attenuates positive demand shocks, which are only partially translated into prices, while negative demand shocks are translated into prices more fully. As Dixit (1999) shows, negative skewness leads firms to delay investment.³ Firms have an incentive to delay investment when large drops in the price of firms' output are more likely. This additional incentive to delay investment, in conjunction with the significantly positive option premia, results in investment delays even longer than predicted by standard, partial-equilibrium real options models.

Because large drops in the price of firms' output result from firms adding capacity they are, *ipso facto*, most likely when firms choose to develop. In fact, firms always add capacity expecting prices to fall. That is, while capacity added immediately prior to large price drops might look, *ex post*, like overbuilding, its development was in fact *ex ante* optimal. We construct forward curves for the price of firms' output to explore the implications of negative skewness, and its business cycle dependency, in greater detail.

We also derive closed-form expressions for firm value as a function of the price of firms' output and current aggregate industry capacity. This allows us to make predictions about stock returns. Stock returns are a combination of the returns to firms' ongoing projects and their growth options. Away from historic highs in the price of firms' output, aggregate capacity is unlikely to increase, so the return to firms' output exhibits little skew. This translates into positive skew in stock returns because the option component of firm value is convex in the price of firm output. When the price of firms' output is high, the effect of increasing aggregate capacity dominates and stock returns exhibit negative skew. This provides a hitherto untested empirical prediction: skewness in stock returns should vary over the business cycle. Stock returns should be negatively skewed during expansions, but positively skewed in recessions.

Finally, we compare the behavior of small competitive firms to that of a strategic monopolist. Not surprisingly, the monopolist, who extracts monopoly rents unavailable to competitive firms, is more profitable

³ Leahy (1993) also notes that a competitive firm's free entry threshold can resemble a monopolist's option-value threshold.

than competitive firms. More interestingly, the fact that the monopolist internalizes the cost that new capacity imposes on ongoing projects results in the monopolist redeveloping projects less frequently, but more intensely. This is expressed in greater heterogeneity in the monopolist economy: the ratio of the most highly developed project to the least developed project in the monopolist economy exceeds that in the competitive economy. We also consider the welfare implications of industrial organization, and find that the lost consumer surplus due to the monopolist restricting supply is partially offset by supply efficiencies. The limited supply the monopolist provides is supplied more efficiently than that provided by competitive firms, which are driven by competitive pressures to redevelop too frequently at inefficiently low capital intensities. This inefficiency is especially pronounced when the price elasticity for the industry's output is high.

The remainder of the article is organized as follows. Section 1 introduces the basic model. Section 2 demonstrates the equilibrium strategy. The section begins by discussing the intuition and the general form of the strategy before characterizing the strategy explicitly. Section 3 discusses implications of the analysis, including properties of the equilibrium price process. Section 4 compares the behavior of competitive agents to both the optimal behavior in the standard, partial-equilibrium model and to the optimal behavior of a strategic monopolist. Section 5 concludes.

1. The Model

In this model we consider a continuum of competitive firms. The defining characteristic of a firm is ownership of productive capital. A firm's ongoing assets, or installed "capacity," costlessly produce a good (or service) flow. A firm can produce a flow of this nonstorable good in proportion to its capacity. The good may be sold in a competitive market at the market clearing price P_t . The total instantaneous cash flow to a site with capacity q , excluding development costs, is therefore qP_t .

Following the literature we assume that the market clearing price for firms' output satisfies an inverse demand function of a constant elasticity form,

$$P_t = X_t Q_t^{-1/\alpha}, \quad (1)$$

where Q_t is the instantaneous aggregate supply of the good, X_t is a multiplicative demand shock, and α is the price elasticity of demand.⁴ This formulation is equivalent to assuming that prices are set by market clearing and demand is time varying but has constant elasticity with respect

⁴ Because projects produce the good in proportion to their capacities Q_t will be used to denote both the instantaneous aggregate supply of the good and the aggregate capacity to produce the good.

to price. That is, demand is given by

$$D_t = X_t^\alpha P_t^{-\alpha},$$

where X_t^α is stochastic and may be thought of as demand in a world in which the good has unit price. The multiplicative demand shock is assumed to evolve as a geometric Brownian motion. That is,

$$dX_t = \mu X_t dt + \sigma X_t dz_t$$

where μ and σ are constant.

At any time a firm may increase capacity by developing. Development, which may be undertaken repeatedly, entails two costs, the investment cost, which is the direct cost of development, and an opportunity cost.⁵

The investment is the cash outlay required to undertake the new project. For example, when a personal computer manufacturer retools a production line for a new model it incurs costs. Likewise, when the owner of a small tenement in Manhattan decides to redevelop her property, she incurs the direct construction costs of building a new office tower.

The direct cost of development depends on the scale of the undertaking. Continuing the previous examples, it is more expensive to set up a production line to produce a million computers a year than it is to set it up to produce a hundred thousand computers; likewise, it is more expensive to build a 60-story office tower than it is to build a 30-story office tower. We model this cost of investment as Cobb-Douglas with increasing costs-to-scale. That is, the direct cost of developing capacity q^* is $q^{*\gamma}$ where $\gamma > 1$.

Opportunity costs result because the new investment damages the firm's ongoing business. In the examples above, for instance, undertaking the new project entails a significant loss in value of the ongoing assets. When the computer manufacturer introduces the new model, it effectively kills demand for the old model. The manufacturer must consider this lost revenue in her investment decision. The owner of the Manhattan tenement is in much the same position. Before putting up the office tower she must first raze the tenement, foregoing future rents.

The opportunity cost to undertaking investment is largely independent of the scale of the undertaking. The computer manufacturer kills demand for the old model whether it produces a hundred thousand units or a million units of the new model, and the tenement owner must raze the existing building regardless of the size of the new office tower.

⁵ While our focus is on the role of competition, allowing for firms that incur opportunity as well as investment costs to increase capacity repeatedly is itself a departure from the standard real options literature, with important implications for optimal investment choice. Retaining development rights leads to higher option values, and to firms developing sooner, but to lower capacities, than they would if they were only able to develop once. See, for example, Williams (1997).

The opportunity cost is modeled as a fractional loss of the value of projects currently in place. For convenience we assume that the fractional loss from adjustment is one, that is, that development entails abandonment of the ongoing project, which simplifies the analysis, but with modification the analysis presented in this article applies to any other choice.

We assume, for the sake of simplicity, that cash flows are valued in a risk-neutral framework, discounted at a constant risk-free rate r . Firms are then priced at the expected value of future revenues, less investment costs, all discounted appropriately for the time value of money. Firms are assumed to maximize value. That is, firms choose their investment strategies to maximize the current expected value of all future cash flows, including development costs, discounted appropriately. We will assume $r > \frac{\alpha(\gamma-1)}{1+\alpha(\gamma-1)}(\mu - \frac{1}{1+\alpha(\gamma-1)}\frac{\sigma^2}{2})$, a restriction that guarantees firm values are finite.

Finally, we would like to capture the fact that firms vary greatly in scope and size. We model heterogeneity in this dimension. The initial heterogeneity is a sort of natural variation in firm size, with firms initially distributed uniformly with respect to log-capacity between the smallest and largest firms in the industry.⁶ That is, the probability of randomly drawing a particular firm of a given size out of the initial distribution is inversely proportional to the firm's size. While we impose this initial distribution on firms, thereafter each firm's size evolves endogenously as a result of its equilibrium investment decisions.

2. Determination of the Equilibrium Strategy

Firms must account for the actions of their competitors when determining their optimal investment strategies. Firms produce a good that they sell in a competitive market at the market clearing price, which is determined by supply and demand. Demand is exogenous but supply is endogenous, resulting from firms' investment decisions.⁷ Firms consequently must invest accounting for the investment strategy of other firms in the industry and the impact of other firms' investment decisions on prices.

The equilibrium concept employed in this article is competitive equilibrium with rational expectations. An equilibrium consists of a consistent (1) price process for the industry good, and (2) set of investment strategies for all firms in the industry. That is, in equilibrium firms' investment decisions are both (1) consistent with the evolution of the price

⁶ Because opportunity costs are proportional to the scale of a firm's ongoing project, this is really an assumption of heterogeneous costs to adjusting capacity. That is, while firms are homogeneous in the cost of investing, they differ with respect to the opportunity cost to undertaking investment.

⁷ Supply is determined purely by firms' investment decisions because in the model firms never idle capacity; as they are small (price takers) and able to produce the good costlessly, they operate at full capacity.

process, and (2) optimal given the price process. As a firm has no incentive to deviate from its investment strategy, readers more comfortable with Nash equilibria may choose to interpret the equilibrium using that concept.

To demonstrate an equilibrium strategy, we must establish two elements. We will first hypothesize an investment strategy and determine the price process that results if all firms adhere to this hypothesized strategy. We will then show that, given the resultant price process, the hypothesized strategy is optimal. That is, we will show that the hypothesized strategy is a fixed point of the mapping from strategies to strategies that results from (i) imposing investment behavior on firms, which maps strategies to price processes, composed with (ii) the unconstrained firm's optimal response, which maps price processes to strategies.

While the goal here is only to demonstrate an equilibrium strategy, we will also establish a limited uniqueness. In particular, the equilibrium presented will be the unique equilibrium for which the evolution of the log-price process depends on the price level P_t only through its relation to the historical maximum $\overline{P}_t$, that is, for which $d \ln P_t$ is a homogeneous, degree-zero function of P_t and $\overline{P}_t$.

2.1 The price of firms' output

We now turn our attention to the first element of establishing an equilibrium strategy, hypothesizing a strategy and determining the resultant price process.

The Strategy Hypothesis . *Suppose that both*

1. *firms' log-capacities are initially distributed uniformly between $\ln q_0$ and $\ln \kappa q_0$, where $\kappa > 1$ satisfies*

$$(\gamma\eta - \gamma - \eta)\kappa^{\gamma\eta - \gamma - \eta + 1} - (\gamma\eta - \gamma)\kappa^{\gamma\eta - \gamma - \eta} + \eta = 0 \quad (2)$$

with $\eta = (1 + \frac{1}{\alpha(\gamma-1)})\beta$ and $\beta = \sqrt{\left(\frac{\mu}{\sigma^2} - \frac{1}{2}\right)^2 + \frac{2r}{\sigma^2}} - \left(\frac{\mu}{\sigma^2} - \frac{1}{2}\right)$, and

2. *the initial price of firms' output is "not too high," satisfying $P_0 \leq q_0^{\gamma-1} P_1^*$ where*

$$P_1^* = \frac{\beta(r-\mu)}{\beta-1} \frac{\kappa^\gamma}{\kappa-1}. \quad (3)$$

Then a firm with existing capacity q optimally redevelops to capacity $q_q^ \equiv \kappa q$ when prices first reach $P_q^* \equiv q^{\gamma-1} P_1^*$.*

The parameter κ , essentially a measure of heterogeneity in the economy, will be explored in greater detail in Section 3.1, after we have established that the hypothesized strategy describes an equilibrium. The log-uniform distribution of firms' capacities is the one in which every firm faces the same degree of competition from similar opportunity cost firms when investing

in new capacity, and consequently this distribution is stable because no firm has an incentive to differentiate itself from similar firms through its investment decisions in an effort to reduce future investment competition.

The multiplicative investment rule of the Strategy Hypothesis preserves the log-uniform cross-sectional distribution of firm capacities, and the smallest firm's capacity is then a sufficient statistic for aggregate capacity. Under the hypothesized strategy, the smallest firm's capacity is $q_t = (\bar{P}_t / P_1^*)^{\frac{1}{\gamma-1}}$. Taken together, these facts allow us to calculate the evolution of aggregate capacity under the hypothesized strategy, which is given in the following proposition. The proofs of all propositions are, in an effort to avoid excessive digression, left for the Appendix.

Proposition 1. *Suppose the conditions of the Strategy Hypothesis hold, and that firms follow the hypothesized strategy. Then aggregate capacity evolves according to*

$$Q_t = \left(\frac{\bar{P}_t}{P_0} \right)^{\frac{1}{\gamma-1}} Q_0 \quad (4)$$

where $\bar{P}_t$, the “historical price maximum up to time t ,” is the greater of $P_{q_0}^*$ and the maximum of P_s on the interval $[0, t]$, and $Q_0 = \left(\frac{\kappa-1}{\ln \kappa} \right) q_0$.

Proposition 1, in conjunction with the inverse pricing rule, $P_t = X_t Q_t^{-1/\alpha}$, allows us to calculate prices explicitly as a function of the multiplicative demand multiplier.

Proposition 2. *Suppose the conditions of the Strategy Hypothesis hold, and that firms follow the hypothesized strategy. Then the price of firms' output evolves according to*

$$\ln P_t = \ln P_0 + \ln \left(\frac{X_t}{X_0} \right) - \frac{1}{1+\alpha(\gamma-1)} \ln \left(\frac{\bar{X}_t}{X_0} \right) \quad (5)$$

where $\bar{X}_t$, the “historical demand maximum up to time t ,” is the greater of $X_{q_0}^* \equiv P_{q_0}^* Q_0^{1/\alpha}$ and the maximum of X_s on the interval $[0, t]$.

The first term on the right hand side of Equation (5) calibrates the level of prices at time zero. The second term expresses the direct effect of demand on prices, and is directly proportional to the current level of the multiplicative demand shock. The third term expresses the effect of supply, as aggregate capacity is related to the historical maximum of price, and consequently demand. Not surprisingly, this supply effect is greater when the cost-to-scale of building is low, or when demand is inelastic with respect to prices.

This price process may be described as a “partially reflected” geometric Brownian motion. Capacity is fixed below historic highs in the price of

firms' output, so the instantaneous evolution of the price process is the same as the evolution of the multiplicative demand shock. Below price highs the only term that changes in the right hand side of Equation (5) in response to a demand shock is $\ln X_t$, so prices change in exactly the same way as demand. At historic highs, however, positive demand shocks result in a supply response that mitigates the effect on prices, and demand shocks are translated into upward movements in the log-price process attenuated by the factor $\frac{\alpha(\gamma-1)}{1+\alpha(\gamma-1)}$.

In Figure 1, this partially reflected geometric Brownian equilibrium price process is shown (middle path) for a particular realization of the evolution of the multiplicative demand shock. Also plotted, for the same realization of the demand shock and starting at the same initial price level, are the standard geometric Brownian motion typical of partial-equilibrium models (top path), and a reflected geometric Brownian motion typical of linear, incremental investment models (bottom path, reflecting barrier at one). The "degree" of reflection in the equilibrium price process depends on the cost-to-scale of adding new capacity, and on the elasticity of prices with respect to supply.

We will explore the properties of the price process in greater detail in Section 3.2, after we have established that it is indeed the equilibrium price process.

2.2 A firm's optimal investment strategy

Now we consider an arbitrary firm with existing capacity q that takes the price process from Proposition 2 as given. Each firm chooses an investment strategy to maximize the expected value of all future cash flows, accounting

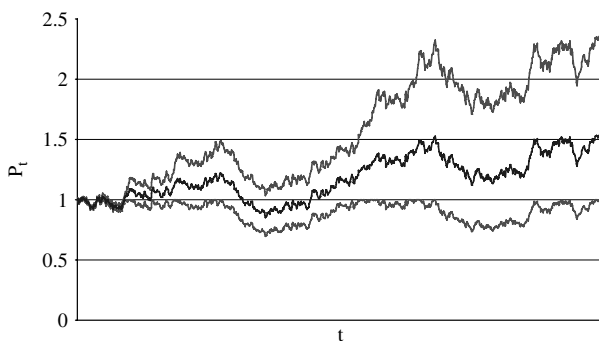


Figure 1
Price Paths

The top path shows a standard, partial-equilibrium, geometric Brownian price process. The bottom path follows reflected geometric Brownian motion. The middle path shows the equilibrium "partially reflected" geometric Brownian price process. The degree of reflection depends on the cost-to-scale of adding capacity and the price elasticity of supply. All three paths are driven by the same realization of the underlying demand process.

for development costs. Adding capacity entails adjustment costs, so firms will not find it optimal to adjust capacity continuously. A firm optimally determines discrete times τ_i at which to increase to capacities q_i , for $i = 1, 2, \dots, \infty$. Letting $V(q, P_t, \bar{P}_t)$ denote the value of a firm with capacity q when the price of firms' output is P_t and the all-time high in the price is $\bar{P}_t$, the value of a firm may then be written as

$$V(q, P_t, \bar{P}_t) = \max_{\{(\tau_i, q_i)\}_{i=1}^{\infty}} E_t \left[\int_t^{\infty} e^{-r(s-t)} q_s P_s ds - \sum_{i=1}^{\infty} e^{-r(\tau_i-t)} q_i^\gamma \right] \quad (6)$$

where $q_s \equiv q_i$ if $s \in [\tau_i, \tau_{i+1})$, and $E_t[\cdot] \equiv E[\cdot | P_t, \bar{P}_t]$ is the expectation conditional on time- t information.

Following the standard methodology, the first step in determining a firm's optimal investment strategy is to decompose firm value into two pieces, the expected value of cash flows received prior to the time at which the next development occurs, and the current discounted expected value of the firm at that time. That is, to write firm value as

$$V(q, P_t, \bar{P}_t) = E_t \left[\int_t^{\tau_1} e^{-r(s-t)} q P_s ds + e^{-r(\tau_1-t)} W(q, P_q^*) \right], \quad (7)$$

where $P_q^* \equiv P_{\tau_1}$ denotes the price level at which the firm next optimally increases capacity, and $W(q, P_q^*) \equiv V(q, P_q^*, P_q^*)$, which we will refer to as the "value-at-max" function, is shorthand for the value of a firm with capacity q when the price of firms' output is at the historical maximum P_q^* . We have implicitly used the fact that firms will only choose to develop at new price maxima because the price process is Markovian below the maxima.

Because firms will only choose to add new capacity at price maxima, it will be sufficient to restrict our attention to these times, in which case we may again decompose firm value into two pieces: the expected value of cash flows received up until the date of the next development, and the present expected value of the firm at that date. That is, we can write firm value as

$$W(q, P_t) = E_t \left[\int_t^{\tau_1} e^{-r(s-t)} q P_s ds \right] + E_t \left[e^{-r(\tau_1-t)} \right] W(q, P_q^*) \quad (8)$$

where we can take $W(q, P_q^*)$ outside the expectation because P_q^* is non-stochastic.

Explicit calculation of the right hand side of Equation (8) is complicated by the fact that the price of firms' output is not an *Itô* process, which prevents direct application of the standard, continuous time machinery. Nevertheless, standard techniques may be applied indirectly, yielding the following proposition.

Proposition 3. *The value of a firm with existing capacity q when the unit price of the firm's output is at an all-time high P_t is given by*

$$W(q, P_t) = \Pi q P_t + \left(\frac{P_t}{P_q^*}\right)^\eta (W(q, P_q^*) - \Pi q P_q^*), \quad (9)$$

where $\Pi = \frac{(\beta-1)\eta}{\beta(\eta-1)}\pi$ and $\eta = \left(1 + \frac{1}{\alpha(\gamma-1)}\right)\beta$.

The first term on the right hand side of Equation (9) is the value of the future cash flows to the firm's existing assets, or the firm's "intrinsic value." The second term, the value of the firm's future development opportunities or "option value," is the firm value in excess of the intrinsic value of assets in place on the date of redevelopment multiplied by the state price for the development date. Currently both firm value on the date of redevelopment and the date at which this development is optimally undertaken are unknown, and depend on the development strategy employed by the firm.

Explicitly determining these values, and the firm's optimal investment strategy, may be accomplished using standard techniques. The functional form of the dependency of firm value on the output price can be determined by inspection of Equation (9). Letting $a_q = (W(q, P_q^*) - \Pi q P_q^*) / P_q^{*\eta}$ we have that

$$W(q, P) = \Pi q P + a_q P^\eta. \quad (10)$$

This, in conjunction with the fact that a firm chooses its development strategy to maximize its value, defines a free boundary problem. The firm's optimal strategy solves this problem, and may be determined using standard procedures.

Optimality and feasibility concerns require value matching and smooth pasting conditions hold at the time of development, i.e., that

$$W^1(q, P_q^*) = W^2(q_q^*, P_q^*) - q_q^{*\gamma} \quad (11)$$

$$W_P^1(q, P_q^*) = W_P^2(q_q^*, P_q^*) \quad (12)$$

where q_q^* denotes the optimal capacity to which a firm with existing capacity q develops when prices reach P_q^* , and W^1 and W^2 are the value-at-max functions before and after development, respectively.

Solving these explicitly is complicated by the fact that firm value is a "compound option." Development rights are retained at exercise, so post development firm value still includes a component due to the right to develop further in the future. The fact that firms retain development rights, however, itself implies a boundary condition that may be used to determine the optimal development strategy explicitly, the results of which are given in the following proposition.

Proposition 4. *Suppose the price of firms' output follows the partially reflected geometric Brownian process given by Equation (5), that is, that $P_t = P_0 \left(\frac{X_t}{X_0} \right) \left(\frac{\bar{X}_t}{\bar{X}_0} \right)^{\frac{-1}{1+\alpha(\gamma-1)}}$ where $d \ln X_t = \mu dt + \sigma dz_t$. Then the optimal development strategy is the Strategy Hypothesis.*

Taken with Proposition 2, Proposition 4 establishes an equilibrium investment strategy. Proposition 2 shows that a partially reflected geometric Brownian price process is consistent with the Strategy Hypothesis. Proposition 4 shows that the optimal investment behavior given the partially reflected geometric Brownian price process is the Strategy Hypothesis. The Strategy Hypothesis is, therefore, an equilibrium strategy.

While we were primarily concerned with identifying an equilibrium, the methodology employed does imply a limited uniqueness. The nature of this limited uniqueness is clarified by introducing the following concept. Think of the ratio $P_t/\bar{P}_t$ as a proxy for the business cycle. When this ratio is close to 1, near-term increases in aggregate industry capacity are likely, and we identify these times with expansions; when this ratio is significantly less than 1, near-term increases in aggregate industry capacity are unlikely, and we identify these times with recessions. Using this concept, the equilibrium presented in this article is the unique equilibrium for which the evolution of log-prices depends on the current level only through the business cycle proxy. This is formalized in the following proposition.

Proposition 5. *The Strategy Hypothesis is the unique equilibrium strategy for which $d \ln P_t$ is a homogeneous, degree-zero function of P_t and $\bar{P}_t$.*

It is also useful to note that we can rule out symmetric equilibrium, the type of equilibrium prevalent in the literature. That is, there is no equilibrium in which homogeneous agents make identical investment decisions. Even if firms are *ex ante* identical, such behavior is *always* suboptimal. An individual firm will invariably find it profitable to deviate from the behavior of other similar firms, and heterogeneity consequently arises endogenously. This is formalized in the following proposition.

Proposition 6. *There is no symmetric equilibrium.*

The intuition behind the preceding result, which can easily be generalized to exclude the existence of any equilibrium in which a non-zero mass of identical agents undertakes simultaneous investment (*i.e.*, behave identically), is as follows. Because simultaneous investment by a positive mass of firms results in a discrete upward jump in aggregate capacity, and thus a discrete downward jump in the price of firms' output, any individual firm prefers either to invest earlier, to capture revenues associated with

the relatively high prices before other firms invest, or to invest later, after prices have recovered. Because no firm finds it optimal to invest when a positive mass of firms are investing, identical firms follow a mixed strategy. Each firm is indifferent between investing early, at relatively low prices and with near-term growth in the output price reduced by the mass of firms still yet to invest, investing late, foregoing intermediate revenues of new production in exchange for investing on a larger scale in better business conditions, and investing at intermediate times, which mix the advantages and disadvantages of the extremes. As a result, mass bleeds out of atoms in the distribution of firm capacities over the course of an investment cycle, and the resultant distribution is smoother than the initial distribution. That is, because firms prefer to invest when they expect strong revenue growth, firms facing excessive competition (i.e., too many firms with similarly low opportunity costs) differentiate themselves through their investment decisions, thereby decreasing the degree of competition they face when making future investment decisions.

While it is relatively easy to prove that the production technology, which entails proportional adjustment costs, cannot support the type of equilibrium most popular in the literature, that is, a symmetric equilibrium, it is more difficult to address the stability of the hypothesized strategy. Here we must satisfy ourselves with heuristic arguments that suggest that the hypothesized equilibrium is robust, and moreover that it is not unreasonable to believe that over successive development cycles the cross-sectional distribution of firm capacities converges to the one given in the strategy hypothesis. The essence of the argument is that in “thick” sections of the distribution of firms’ capacities, firms that build later build more aggressively. A low opportunity cost firm that invests early in this region of the distribution does so knowing that near-term price growth will be strongly attenuated by the relatively large number of other firms with opportunity costs nearly as low, which will also be adding new capacity in the near future.

Later, when a slightly higher opportunity cost firm decides to invest, it does so in a better business environment: the relatively large mass of firms between this firm and the original firm have already invested, so are no longer threatening to invest in new capacity in the near future. This leads the firm that invests later to invest to a slightly higher capacity multiple. The difference between these firms’ log-capacities is consequently greater after development than before. That is, development thins out the capacity distribution in regions where it is too thick. In “thin” sections of the distribution the opposite occurs; firms that build later build less aggressively, so new development tends to thicken these sections of the distribution. The cross-sectional distribution of firms’ capacities given in the Strategy Hypothesis is the “Goldilocks” distribution. The log-uniform distribution of firms’ capacities, with and only with width $\ln \kappa$, is “just

right,” in the sense that it is neither “too thick” nor “too thin” at any point. With this distribution, every firm faces the same degree of competition, and no firm has an incentive to deviate their investment strategy to get to a point in the distribution with fewer similar firms, because no such point exists.

2.3 Firm valuation

Having established that the Strategy Hypothesis is indeed an equilibrium strategy, we may value firms explicitly. The equilibrium valuation of an arbitrary firm is given in the following proposition.

Proposition 7. *The value of a firm with existing capacity q when the price of the firm’s output is P_t and the historical price maximum is $\bar{P}_t \leq P_q^*$ is given by*

$$V(q, P_t, \bar{P}_t) = \pi q P_t - \left(\frac{P_t}{\bar{P}_t}\right)^\beta (\pi - \Pi) q \bar{P}_t + \left(\frac{P_t}{\bar{P}_t}\right)^\beta \left(\frac{\bar{P}_t}{P_q^*}\right)^\eta \frac{\kappa^\gamma q^\gamma}{(\eta-1)(1-\kappa^\gamma + \eta - \gamma\eta)}. \quad (13)$$

The first term in Equation (13) is the value of the cash flows from the current project, ignoring supply effects on prices. The second term corrects for these supply effects, and is negative reflecting the fact that new capacity puts downward pressure on the price of firms’ output. The last term is the option value of the project, which derives from the firm’s ability to increase capacity in the future.

3. Properties of the Equilibrium

Before considering how the investment behavior of a competitive firm differs from the predictions of the standard partial-equilibrium model, or from that of a strategic monopolist, we will consider some properties of the competitive equilibrium in greater detail.

3.1 Equilibrium investment behavior

In the competitive equilibrium, a firm optimally develops to κ times its existing capacity. The optimal timing of this development also depends fundamentally on the parameter κ , as P_1^* is proportional to $\kappa^\gamma / (\kappa - 1)$. The parameter κ , essentially a measure of heterogeneity in the economy, is not arbitrary, but a fixed constant that depends on the primitives of the economy. The following proposition demonstrates how this parameter varies with the economic primitives.

Proposition 8. *The optimal development multiple κ*

1. *is uniquely determined by the economic primitives of the economy $(r, \mu, \sigma, \alpha, \text{ and } \gamma)$,*

2. is strictly decreasing in r , and strictly increasing in μ , σ , and α ,
3. approaches the lower bound $\underline{\kappa} = \frac{\gamma}{\gamma-1}$ as $r - \mu$ gets large, or as α gets small, and
4. approaches the upper bound $\bar{\kappa}$ as η approaches $\frac{\gamma}{\gamma-1}$, where $\bar{\kappa}$ satisfies $\frac{\bar{\kappa}-1}{\ln \bar{\kappa}} = \frac{\gamma}{\gamma-1}$.

These dependencies are depicted graphically in Figure 2, which shows the optimal development multiple as a function of interest rates, the growth rate and the volatility of the multiplicative demand multiple, and the cost-to-scale of developing capacity.

3.2 The equilibrium price process

The model predicts that the price of firms' output will be negatively skewed, resulting from endogenous increases in capacity. Because capacity is added at price maxima, this skewness is especially pronounced when prices are high. In fact, at maxima in the price of firms' output, we always expect prices to drop in the short term, and for some parameterizations to remain

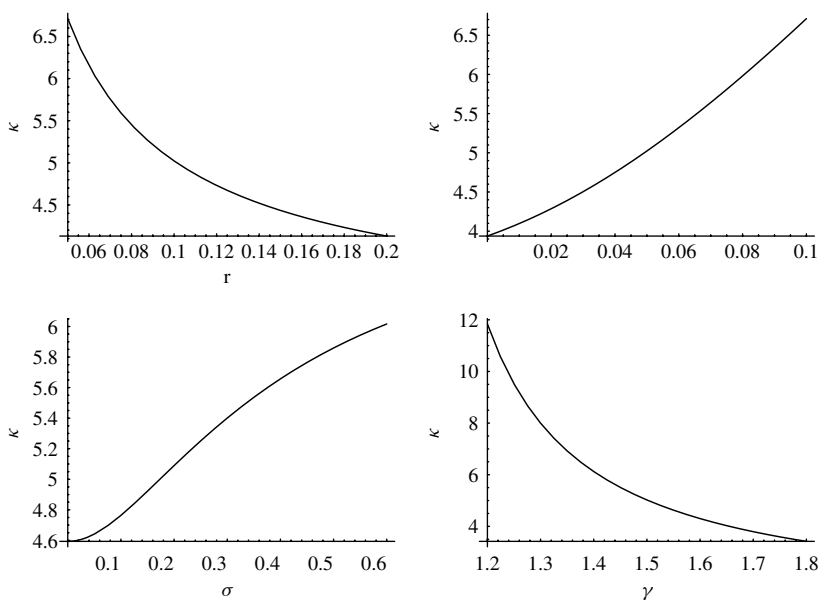


Figure 2
Comparative Statics for the Optimal Development Multiple

The figure shows the optimal development density multiple, κ , as a function of the economy's primitives. The optimal development multiple is (1) decreasing in the interest rate (top right), as higher rates decrease the marginal value of development, (2) increasing in demand growth for the industry good (top right), as higher growth rates reduce the likelihood of weak future demand encouraging greater investment; (3) increasing in uncertainty regarding future demand (bottom left), as higher uncertainty incentivizes firms to delay investment, and (4) decreasing, not surprisingly, in the cost-to-scale of development (bottom right). Default parameter values are $r = .06$, $\mu = .04$, $\sigma = .18$, $\alpha = 1$ and $\gamma = 2$.

low far into the future. As firms only add capacity when the price of their output is high, this means that firms always develop expecting prices to drop.

The negatively skewed equilibrium return to firms' output is itself a major prediction of the model and provides testable implications. Because we characterize the equilibrium price process analytically, we can make specific predictions regarding properties of the skewness. The model predicts the degree of skewness to be a function of both industry cost structure and industry history. Cross-sectionally, we would expect more pronounced price skewness in industries where the supply elasticity of demand is high. That is, prices should be more skewed in industries where the costs of adding capacity are low. We would also expect more pronounced price skewness when the demand for firms' output is elastic.

In the time series, we expect to see less price skewness at short horizons. The model predicts supply to be more responsive when prices are near historic highs. At lower prices, when firms are less likely to add capacity, skewness should be less pronounced, and return skewness should tend to zero as the horizon becomes very short. At price maxima, on the other hand, skewness in prices should be especially pronounced.

Forward prices provide the most convenient method for studying the price process in greater detail. The forward price in this economy, in which investors are risk-neutral, is the unbiased estimate of the future spot price, $F_{0,t} = E_0[P_t]$. The closed-form expression for the forward price of firms' output is provided in the next proposition. It is somewhat complicated, but studying the behavior of some of its asymptotic properties proves illuminating.

Proposition 9. *The equilibrium t -ahead forward price of the good (i.e., the expected future spot price) is given by*

$$F_{0,t} = e^{\mu t} P_0 G_{\mu,\sigma} \left(\frac{\bar{P}_0}{P_0}, \frac{1}{1+\alpha(\gamma-1)}, t \right) \quad (14)$$

where

$$\begin{aligned} G_{\mu,\sigma}(x, \Omega, t) = & N \left(\frac{\ln x - \left(\mu + \frac{\sigma^2}{2} \right) t}{\sigma \sqrt{t}} \right) + \left(\frac{\Omega}{1 + \frac{2\mu}{\sigma^2} - \Omega} \right) N \left(\frac{-\ln x - \left(\mu + \frac{\sigma^2}{2} \right) t}{\sigma \sqrt{t}} \right) x^{1 + \frac{2\mu}{\sigma^2}} \\ & + \left(1 - \frac{\Omega}{1 + \frac{2\mu}{\sigma^2} - \Omega} \right) e^{-\Omega \left(\mu + (1-\Omega) \frac{\sigma^2}{2} \right) t} N \left(\frac{-\ln x + \left(\mu + (1-2\Omega) \frac{\sigma^2}{2} \right) t}{\sigma \sqrt{t}} \right) x^{\Omega}. \end{aligned}$$

Figure 3, which shows forward prices at horizons out to two years, demonstrates the importance of the price to price-high ratio on forward prices. The figure plots three possible relations between prices and the historical price high: (1) spot prices at the high; (2) spot prices one-sixth

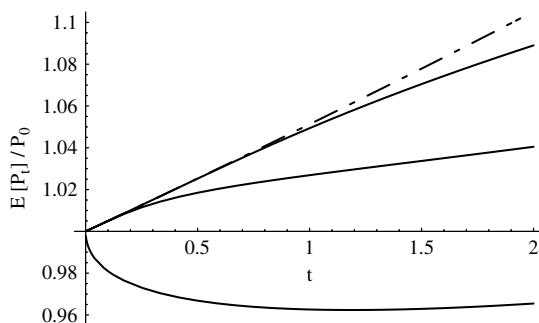


Figure 3
Equilibrium Forward Prices

The forward price is the unbiased estimator of the future spot price. The bottom curve shows forward prices when the spot price is at the historical maximum, $P_0 = \bar{P}_0$. The middle curve shows forward prices when the spot price is 20% below the historical maximum, $P_0 = 0.8\bar{P}_0$. The top curve shows forward prices when the spot price is 40% below the historical maximum, $P_0 = 0.6\bar{P}_0$. The dashed line shows forward prices assuming fixed aggregate capacity (i.e., assuming no supply effects). Parameter values are $\mu = .04$, $\sigma = .18$, $\alpha = 1$ and $\gamma = 2$.

below the high; and (3) spot prices one-third below the high. In the figure, the cost-to-scale of development is quadratic.

The basic shapes of the forward price curves in Figure 3 are similar to those observed in commodity markets. When the current price of firms' output is at or near historic highs, increases in aggregate capacity are likely. At these times, forward prices are downward sloping, or "backwarddated," with the future price declining with time-to-delivery. When the current price of firms' output is significantly below historic highs, then increases in aggregate capacity are unlikely. At these times, forward prices are upward sloping, or "in contango," with prices increasing with time-to-delivery. Not surprisingly, the model predicts more frequent and more pronounced backwardation in industries that produce goods for which demand is inelastic with respect to prices, and in industries where the cost-to-scale of adding new capacity is low. When demand is inelastic, new supply elicits a large price response, directly attenuating price growth. When the cost of new capacity is low, rising prices elicit a large supply response, which in turn attenuates price growth.

We can also use forward prices to consider the average long-run growth in spot prices, obtained by evaluating $\ln F_{0,t}/t$ in the limit as time-to-delivery gets large. Doing so, in the interesting case when μ is not too small, we find that the average rate of expected price growth is $\frac{\rho}{\eta} \left(\mu - (1 - \frac{\rho}{\eta}) \frac{\sigma^2}{2} \right)$.

Finally, negative skewness in the price of firms' output has implications for the equilibrium evolution of *stock* prices. The negative skewness in output prices is translated into stock prices, but only partially. Firm value is convex in the price of the firm's output, and this tends to skew stock prices *positively*. When the price of firms' output is low, output prices

are essentially unskewed, and the effect of convexity dominates. At these times, stock prices are positively skewed. When the price of firms' output is high, however, the effect of the strong negative skewness in the price of firms' output dominates. At these times, stock prices are also negatively skewed. This provides a hitherto untested empirical prediction: skewness in stock returns should vary over the business cycle. Stock returns should be negatively skewed during expansions, but positively skewed in recessions.⁸

Proposition 10. *At maxima in the price of firms' output (i.e., in industry expansions) stock returns are negatively skewed at short horizons. Away from maxima in the price of firms' output (i.e., in industry recessions) stock returns are positively skewed at sufficiently short horizons.*

4. Comparisons to Other Economies

We now consider how the competitive equilibrium compares with two important alternatives: (1) the standard partial-equilibrium model, and (2) the monopoly economy. We are concerned with the standard, partial-equilibrium model because this is the model that has largely informed investors' intuition regarding real options. We are concerned with the monopoly economy because it reveals welfare implications regarding industry organization.

Our comparison to the standard, partial-equilibrium model will focus on how the investment decisions of competitive firms differ from the predictions of the standard partial-equilibrium model. That is, we consider how a competitive firm's behavior compares to the expectations of an econometrician armed with the standard model and time-series data on the observable economic variable, prices. We also consider the implications of this behavior on real option values.

Our comparison to the monopoly economy focuses on how the behavior of competitive firms differs from what we would observe if the industry consisted of a single monopolistic firm. That is, we consider how the monopolist's behavior deviates from that of competitive agents, given the same evolution of the primitive economic demand variable, the multiplicative demand multiplier. We also consider the welfare implications of these differences.

4.1 Comparison to the standard model

The standard, partial-equilibrium model assumes that the price of firms' output evolves as geometric Brownian motion, and is equivalent to the model presented in this article assuming firms' investment does not affect

⁸ Exit will also tend to induce positive skewness in recessions. Refinements to Proposition 10 are needed, quantifying the business cycle dependence of the term structure of skewness, to distinguish between these two possible sources of positive skew.

aggregate capacity, an assumption often justified by the fact that an individual firm's investment decisions do not significantly impact aggregate supply if firms are small. But even when firms are price takers individually, their investment decisions still determine the evolution of supply in aggregate. Moreover, firms make price contingent, and thereby highly correlated, investment decisions. So even if an individual firm's decision to increase capacity does not significantly impact aggregate capacity, its decision to increase capacity is nevertheless strongly correlated with increasing aggregate capacity, a fact effectively ignored by the standard model.

The analysis here demonstrates that accounting for the price impact of development actually leads firms to develop *later* than predicted by the standard model. That is, in equilibrium, firms delay capital improvements longer than the standard, partial-equilibrium analysis prescribes. It also shows that competition does not erode option premia: after properly accounting for the impact of competition, future development rights can contribute just as large a fraction of overall project value as they do in the standard analysis that ignores competition.

4.1.1 The investment strategy. An econometrician who calibrates a real options model under the standard assumption in the literature, that the price of firms' output follows a geometric Brownian motion, will be surprised to find firms delaying irreversible investment longer than his model predicts is optimal. This is formalized in the following proposition.

Proposition 11. *In equilibrium a firm redevelops a project to the same capacity as it would if prices were geometric Brownian with the equilibrium long-run average growth and variance, but does so later (i.e., at higher prices). That is, letting*

$$\tilde{\mu} = \lim_{t \rightarrow \infty} \frac{E_0[P_t/P_0]}{t}$$

$$\tilde{\sigma} = \lim_{t \rightarrow \infty} \frac{Var[P_t/P_0]}{t}$$

denote the long-run average growth and variance of the equilibrium price process, then

$$\kappa \equiv \tilde{\kappa}_{p.e.} \quad (15)$$

$$P_1^* \equiv \frac{\beta(\eta-1)}{(\beta-1)\eta} \tilde{P}_1^* > \tilde{P}_{1,p.e.}^* \quad (16)$$

where $\tilde{\kappa}_{p.e.}$ and $\tilde{P}_{1,p.e.}^$ denote the optimal development multiple and price level trigger in a partial-equilibrium model in which the price of firms' output follows a geometric Brownian process with growth rate $\tilde{\mu}$ and volatility $\tilde{\sigma}$.*

Firms delay investment longer because, in addition to the delay related to option effects, the equilibrium return to firms' output is negatively skewed.⁹ Option value, and the associated option related investment delay, derive from the ability to avoid some investments that would have been mistakes *ex post*. At the zero NPV threshold an investor may do two things: invest or delay. The choice to delay roughly reflects the fact that, at the current price level, the *ex post* benefit to having not invested if the economic environment takes a turn for the worse exceeds the *ex post* benefit to having invested earlier if the environment takes a turn for the better. Investment occurs when prices are just high enough that the loss-avoidance benefits to delaying balance the gains from not delaying. Asymmetric downside risk exceeding upside potential induces firms to delay option exercise longer. The oversized downside risk results in the potential *ex post* benefits to delaying investment balancing the *ex post* benefit to not delaying investment at a higher price level.

4.1.2 The value of the option to delay investment. The analysis also shows that competition leaves option premia unmitigated. Competition erodes option values, but also erodes the value of the cash flows from assets in place. As a result, the premium, or percent of total option value attributable to future development opportunities, is undiminished relative to an economy in which prices follow geometric Brownian motion and have the same average long-run price growth and variance. That is, option premia are unmitigated relative to the partial-equilibrium analysis that ignores competition. To formalize this, note that the first term in the value-at-max function, $\Pi q P$, is the value of cash flows from assets in place, or "intrinsic value," of the project. The residual, $W(q, P) - \Pi q P$, derives from the ability to add capacity in the future, and represents the "option value" of the project. The fact that option premia remain significant in equilibrium is given in the following proposition.

Proposition 12. *The ratio of option value to intrinsic value at any price maxima is given by*

$$\frac{W(q, P) - \Pi q P}{\Pi q P} = \frac{\kappa - 1}{\eta (1 - \kappa^{\gamma + \eta - \gamma \eta})} \left(\frac{P}{P_q^*} \right)^{\eta - 1}. \quad (17)$$

This ratio is strictly positive, and for some parameterizations exceeds one, in which case the majority of a project's value resides in the development option.

To show that option premia are unmitigated relative to the partial-equilibrium analysis we will compare the expectations of an econometrician

⁹ "Return to firms' output" here means log-change in output price over elapsed time.

armed with price data and the standard, partial-equilibrium model to the actual results observed in the equilibrium economy. If one calculates the option premia assuming the correct annualized long-run average price growth and variance, $\tilde{\mu}$ and $\tilde{\sigma}$ as in Proposition 11, but incorrectly assuming geometric Brownian prices (or, equivalently, that prices are completely inelastic with respect to supply), then $\tilde{\kappa} = \kappa$, $\tilde{\mu} = \frac{\beta}{\eta} \left(\mu - (1 - \frac{\beta}{\eta}) \frac{\sigma^2}{2} \right)$ and $\tilde{\sigma} = \frac{\beta}{\eta} \sigma$, so $\tilde{\eta} = \tilde{\beta} = \sqrt{(\frac{\tilde{\mu}}{\tilde{\sigma}^2} - \frac{1}{2})^2 + \frac{2r}{\tilde{\sigma}^2}} - (\frac{\tilde{\mu}}{\tilde{\sigma}^2} - \frac{1}{2}) = \eta$, and Equation (17) becomes

$$\frac{\tilde{\kappa} - 1}{\tilde{\eta} \left(1 - \tilde{\kappa}^{\gamma + \tilde{\eta} - \gamma \tilde{\eta}} \right)} \left(\frac{\tilde{P}}{\tilde{P}_q^*} \right)^{\tilde{\eta} - 1} = \frac{\kappa - 1}{\eta \left(1 - \kappa^{\gamma + \eta - \gamma \eta} \right)} \left(\frac{\tilde{P}}{\tilde{P}_q^*} \right)^{\eta - 1}, \quad (18)$$

which immediately implies the following corollary.

Corollary 1. *The ratio of option value to intrinsic value, as a function of the ratio of output price to exercise price, is the same in the equilibrium economy as it is in the partial-equilibrium economy with the same long-run average price growth and variance.*

4.2 Comparison to the strategic monopolist

We now turn our attention to the strategic monopolist. We consider how the monopolist's behavior differs from that of competitive firms. We also examine the consequent welfare implications.¹⁰

4.2.1 The monopolist's optimal investment strategy. The strategy can be characterized by specifying the capacity to which the monopolist will redevelop a site with any existing capacity q , which we will denote $\bar{q}_q^*$, and the price level at which the development will occur, which we will denote M_q^* . It is given explicitly in the following proposition.

Proposition 13. *Suppose that both*

1. *the log-capacities of the monopolistic firm's projects are initially distributed uniformly between $\ln q_0$ and $\ln \bar{\kappa} q_0$, where $\bar{\kappa} > 1$ satisfies $\frac{\bar{\kappa} - 1}{\ln \bar{\kappa}} = \frac{\gamma}{\gamma - 1}$, and*
2. *the initial price of the firm's output is "not too high," satisfying $P_0 \leq q_0^{\gamma - 1} M_1^*$ where*

$$M_1^* = \left(\frac{\alpha}{\alpha - 1} \right) \frac{\beta}{(\beta - 1)\pi} \frac{\bar{\kappa}^\gamma}{\bar{\kappa} - 1}. \quad (19)$$

¹⁰ We assume that aggregate revenues increase with supply (i.e., that the price elasticity of the industry good is greater than 1), otherwise the monopolist will never develop new capacity.

Then the monopolist optimally develops a site with existing capacity q to capacity $\bar{q}_q^* \equiv \bar{\kappa}q$ when prices first reach $M_q^* \equiv q^{\gamma-1}M_1^*$.

The monopolist's optimal investment strategy is similar to that of firms in the competitive equilibrium, in that both strategies follow a multiplicative investment rule of (i) redeveloping a site to a fixed multiple of its existing capacity, and (ii) undertaking this development at a multiple of the price level at which the previous development was undertaken. The monopolist's strategy differs, however, in both the intensity and timing of development.

In the absence of the pressures of competition, the monopolist chooses the redevelopment intensity, $\bar{\kappa}$, which minimizes the cost of supplying any new aggregate capacity. That is, new capacity developed by the monopolist is developed as efficiently as possible. Note that $\bar{\kappa}$, which is also a measure of heterogeneity in the monopoly economy, is the upperbound on the competitive development multiple, so the monopoly economy exhibits greater cross-sectional variation in project size.

In the absence of the pressures of competition, the monopolist times her development to maximize the value of *aggregate* capacity. The monopolist internalizes the cost (lost revenues) that new development imposes on existing capacity, and therefore delays adding new capacity beyond the time at which competitive firms would do so. In fact, not only is the development threshold higher, but it is also sufficiently higher that the monopolist rations aggregate supply, forcing prices higher than in the competitive economy. This is formalized in the next proposition, where Q_t denotes the aggregate supply that would prevail in the competitive economy and $\hat{Q}_t$ denotes the aggregate supply that would prevail in the monopoly economy.

Proposition 14. *The monopolist always rations supply: $\phi \equiv \hat{Q}_t/Q_t < 1$.*

4.2.2 Welfare implications. Monopoly results in both a welfare loss and gain. The welfare loss occurs because the monopolist restricts supply, while the welfare gain results from the fact that the monopolist supplies the good more efficiently than competitive firms.

The instantaneous welfare loss associated with the monopolist's supply rationing is the lost consumer surplus, which we denote S_t . This is given by

$$\begin{aligned} S_t &= \hat{Q}_t (\hat{P}_t - P_t) + \int_{\hat{Q}_t}^{Q_t} \left(\frac{x_t}{Q_t^{1/\alpha}} - P_t \right) dQ \\ &= \frac{1-\phi^{1-1/\alpha}}{\alpha-1} Q_t P_t, \end{aligned} \quad (20)$$

where the integration is performed using the fact that $\hat{Q}_t = \phi Q_t$, so $\hat{P}_t = \phi^{-1/\alpha} P_t$.

The welfare gain associated with monopoly results because the monopolist provides the rationed supply efficiently, whereas competitive pressures drive competitive firms to add capacity, and incur the associated adjustment costs, too frequently.

The *net* impact is an expected discounted deadweight loss due to monopoly, denoted L_t and given in the next proposition. The proposition concerns itself with expansionary times, though a more general formulation is given in the Appendix.

Proposition 15. *The expected discounted deadweight loss from monopoly at price maxima is given by*

$$L_t = \frac{1-\phi^{1-1/\alpha}}{\alpha-1} Q_t P_t \varpi - \left(1 - \frac{\alpha-1}{\alpha} \phi^{1-1/\alpha}\right) \left(\frac{\ln \kappa}{\kappa-1} Q_t\right)^\gamma \Gamma \quad (21)$$

where

$$\begin{aligned} \phi &= \left[\left(\frac{\frac{\kappa-1}{\ln \kappa}}{\frac{\kappa-1}{\ln \kappa}} \right) \left(\frac{P^*}{M_1^*} \right)^{\frac{1}{\gamma-1}} \right]^{\beta/\eta} \\ \varpi &= \frac{\beta-1}{\beta} \frac{(\gamma-1)\eta}{(\gamma\eta-\gamma-\eta)} \pi \\ \Gamma &= \frac{\kappa^\gamma}{(\gamma\eta-\gamma-\eta) \ln \kappa}. \end{aligned}$$

The parameter ϕ measures the degree to which the monopolist restricts supply, ϖ is the dollar value of a unit revenue flow under the aggregate revenue process, and Γ is the expected discounted aggregate cost of all future development in the competitive economy when the smallest project has unit capacity.

As α goes to 1, the monopolist's rationing of supply becomes more severe, and the expected discounted dead weight loss, which asymptotically goes to $\frac{Q_t P_t \varpi}{\alpha-1} - \left(\frac{\ln \kappa}{\kappa-1} Q_t\right)^\gamma \Gamma$, becomes arbitrarily large as P_t becomes large. That is, the dead weight loss associated with the monopoly production of goods for which demand is price-inelastic (e.g., cigarettes or salt) can be arbitrarily high.

As α becomes large the expected discounted deadweight loss to monopoly goes to $(\phi-1)(q_t^m)^\gamma \Gamma < 0$. That is, the deadweight loss can be negative; if the price elasticity of demand is high enough (i.e., if increasing supply hardly impacts prices) then the deadweight loss from the monopolist's rationing of supply is inconsequential, while the efficiency gains to monopoly supply may be significant.

5. Conclusion

We characterize firms' optimal investment strategy explicitly, and derive a closed-form solution for firm value. We show that in the strategic equilibrium, real option premia are significant. As a result, firms delay

investment, choosing optimally not to undertake some positive NPV projects. In fact, properly accounting for the effects of competition results in a negatively skewed equilibrium process for the price of firms' output, which leads firms to delay investment longer than predicted by the standard options analysis. Finally, the equilibrium analysis in this article has implications for equity markets. The model presented here predicts that firm returns should vary over the business cycle. Firm returns should be negatively skewed during expansions, but positively skewed in recessions.

These conclusions differ from those in previous studies of competition on option exercise, which have tended to conclude that competition reduces option premia and leads to earlier investment. Previous studies have considered the effects of oligopolistic competition and have quite naturally, therefore, compared valuation and optimal timing of investment to the case of a hypothetical strategic monopolist. While their conclusions are valid in this context, they should not be used to draw conclusions with respect to the standard analysis, which ignores competition. The standard analysis is not a fundamentally better yardstick than the strategic monopolist. It is, however, the theory that has informed the intuition that is now widely held, and the fact that our methodology allows for comparison to the standard analysis is, therefore, a great advantage.

The results in this article differ from those in earlier articles because we consider *heterogeneous* firms facing opportunity costs to investing. While it has long been known that demand side heterogeneity can reduce competition, in this article we demonstrate that supply side heterogeneity can have the same effect, though for different reasons. Demand side heterogeneity—variable consumer preferences—allows firms to segment a market through horizontal product differentiation, and extract more of the consumer surplus. Supply side heterogeneity—variable opportunity costs—provides a natural ordering to firms' investment decisions, allowing firms to act sometimes as local monopolists and extract more of the consumer surplus. The model predicts a “life cycle” to firms' investment decisions, because firms that have invested recently find it unprofitable to invest again for some time. This prevents them from competing over new opportunities, and significantly limits the role of competition.

Appendix

Proof of Proposition 1. Suppose firms' log-capacities are initially distributed uniformly

between some minimum $\ln q_0^m$ and maximum $\ln q_0^M$, and that each firm follows the strategy

of developing to a multiple $k \equiv q_0^M / q_0^m$ multiplied by its existing capacity q whenever the price of firms' output reaches $(q/q_0^m)^{\gamma-1} \bar{P}_0$.

The aggregate supply process is the integral of individual firms' capacities,

$$Q_t = \int_{q_t^m}^{q_t^M} q d\nu_t(q).$$

The development rule preserves the uniform distribution of log-capacities. That is, at all times $\bar{P}_t$ is a sufficient statistic for aggregate capacity and the distribution of capacities is $(\xi/k)q_t^M$, where $\ln_k \xi$ is distributed uniformly on $(0, 1]$, and

$$\begin{aligned} q_t^M &= q_0^M k^{\ln_k(\bar{P}_t/\bar{P}_0)} \\ &= q_0^M (\bar{P}_t/\bar{P}_0)^{\frac{1}{\gamma-1}}, \end{aligned}$$

so

$$Q_t = q_0^M \left(\frac{\bar{P}_t}{\bar{P}_0} \right)^{\frac{1}{\gamma-1}} \int_0^k \xi d\nu_t(\xi) = \left(\frac{\bar{P}_t}{\bar{P}_0} \right)^{\frac{1}{\gamma-1}} Q_0. \quad (\text{A1})$$

That the aggregate capacity process evolves in a manner *independent* of the width of the distribution of log-capacities is quite remarkable. This independence results from modeling opportunity costs as abandonment of ongoing assets. If we choose to model opportunity costs as some other fractional loss f of the value of assets in place, we would need to replace $\gamma - 1$ with $(1 + \ln_k(1 + \frac{1-f}{k}))^{-1}(\gamma - 1)$. ■

Proof of Proposition 2. The inverse demand function, $P_t = X_t Q_t^{-1/\alpha}$, is valid everywhere, so holds at historic price highs. These correspond to historic highs in the multiplicative demand shock, so $\bar{P}_t = \bar{X}_t Q_t^{-1/\alpha}$. Substituting this for $\bar{P}_t$ in Equation (A1) and simplifying we then have

$$Q_t = \left(\frac{\bar{X}_t}{\bar{X}_0} \right)^{\frac{1}{\gamma-1+\alpha^{-1}}} Q_0.$$

Substituting back into the inverse pricing function yields the proposition. ■

Note that the supply elasticity of the aggregate demand maximum (i.e., the multiplicative shock maximum) is constant, and equal to $(\gamma - 1 + 1/\alpha)^{-1}$. Not surprisingly, supply is more responsive to demand when the cost-to-scale of building is low and when demand is inelastic with respect to prices.

Proof of Proposition 3.

lemma 1. Let Y_t be a geometric Brownian motion with any drift, beginning at 1. Then the value of cash flows proportional to the process and received until the first time the process reaches $\theta > 1$ is given by

$$\frac{1 - \theta^{1-\beta}}{r - \mu}. \quad (\text{A2})$$

Proof of lemma: Let $\tau_\theta \equiv \min\{t > 0 | Y_t = \theta\}$. We want to evaluate, for any μ ,

$$E_0 \left[\int_0^{\tau_\theta} e^{-rt} Y_t dt \right], \quad (\text{A3})$$

Now Y_t is distributed $e^{\sigma(B_t + (\frac{\mu}{\sigma} - \frac{\sigma}{2})t)}$ where B_t is a standard Brownian motion with unit volatility, so the previous integrand may be rewritten, after changing the measure to demean

the Brownian motion and using the standard probability literature result that the joint density for the value and the maximum of a standard Brownian (See, for example, Durrett (1996)) is $v(b, m) = \sqrt{2/\pi} ((2m - b)/t\sqrt{t}) e^{-(2m-b)^2/2t}$, as

$$\int_{t=0}^{\infty} \int_{m=0}^{\frac{\ln(\theta)}{\sigma}} \int_{b=-\infty}^m e^{-rt} e^{\sigma b} \left(\sqrt{\frac{2}{\pi}} \frac{2m-b}{t\sqrt{t}} e^{\frac{-(2m-b)^2}{2t}} \right) e^{\left(\frac{\mu}{\sigma} - \frac{\sigma}{2}\right)b - \left(\frac{\mu}{\sigma} - \frac{\sigma}{2}\right)^2 \frac{t}{2}} db dm dt. \quad (A4)$$

Using the fact, from the Bessel function literature, that

$$\int_{t=0}^{\infty} e^{-\alpha t} \sqrt{\frac{2}{\pi}} \frac{2m-b}{t\sqrt{t}} e^{\frac{-(2m-b)^2}{2t}} dt = 2e^{-\sqrt{2\alpha}(2m-b)}, \quad (A5)$$

Equation (A4) may be integrated directly, with $\alpha = r + \frac{1}{2} \left(\frac{\mu}{\sigma} - \frac{\sigma}{2} \right)^2$, yielding

$$\begin{aligned} & 2 \int_{m=0}^{\frac{\ln(\theta)}{\sigma}} e^{-2\sqrt{2r + \left(\frac{\mu}{\sigma} - \frac{\sigma}{2}\right)^2} m} \left(\int_{b=-\infty}^m e^{\left(\left(\frac{\mu}{\sigma} + \frac{\sigma}{2}\right) + \sqrt{2r + \left(\frac{\mu}{\sigma} - \frac{\sigma}{2}\right)^2}\right)b} db \right) dm \\ &= \frac{2}{\left(\frac{\mu}{\sigma} + \frac{\sigma}{2}\right) + \sqrt{2r + \left(\frac{\mu}{\sigma} - \frac{\sigma}{2}\right)^2}} \int_{m=0}^{\frac{\ln(\theta)}{\sigma}} e^{\left(\left(\frac{\mu}{\sigma} + \frac{\sigma}{2}\right) - \sqrt{2r + \left(\frac{\mu}{\sigma} - \frac{\sigma}{2}\right)^2}\right)m} dm \\ &= \frac{1 - \theta^{1-\beta}}{r - \mu}. \end{aligned} \quad (A6)$$

Alternatively, the lemma may be proved by restricting attention to the case when $\mu < r$, in which case the valuation is standard, and using analytic continuation to argue that the resulting solution is valid for all μ . ■

lemma 2. Suppose Y_t^δ follows partially reflected geometric Brownian motion where $1 - \delta$ denotes the degree of reflection. That is, $Y_t^\delta = \exp(X_t^\delta)$, where $X_t^\delta = \delta \bar{X}_t - (\bar{X}_t - X_t)$, X_t is a drifted Brownian motion and $\bar{X}_t$ is the maximum of the Brownian process up to time t . Then the dollar price of a unit cash flow proportional to the process is given by

$$\pi_\delta = \left(\frac{\beta - 1}{\beta - \delta} \right) \pi, \quad (A7)$$

subject to the parameter restriction $\mu < \frac{r}{\delta} + (1 - \delta) \frac{\sigma^2}{2}$, which ensures π_δ is finite.

Proof of lemma: The dollar price of the unit cash flow is

$$\pi_\delta = \mathbf{E} \left[\int_0^\infty e^{-rt} \left(\frac{Y_t^\delta}{Y_0^\delta} \right) dt \right]. \quad (A8)$$

Letting τ_θ^δ be the stopping time for the first time Y_t^δ hits $\theta > 1$, we then have

$$\begin{aligned} \pi_\delta &= \mathbf{E}_0 \left[\int_0^{\tau_\theta^\delta} e^{-rt} Y_t^\delta dt \right] + Y_{\tau_\theta^\delta}^\delta \mathbf{E}_0 \left[e^{-r\tau_\theta^\delta} \right] \mathbf{E}_0 \left[\int_{\tau_\theta^\delta}^\infty e^{-r(t-\tau_\theta^\delta)} \frac{Y_t^\delta}{Y_{\tau_\theta^\delta}^\delta} dt \right] \\ &= \mathbf{E}_0 \left[\int_0^{\tau_\theta^\delta} e^{-rt} Y_t^\delta dt \right] + \theta \rho_\delta(1, \theta) \pi_\delta, \end{aligned} \quad (A9)$$

where $\rho_\delta(1, \theta) = \rho(1, \theta)^{\frac{1}{\delta}} = \theta^{-(\beta/\delta)}$. Rearranging gives

$$\pi_\delta = \frac{E_0 \left[\int_0^{\tau_\theta^\delta} e^{-rt} Y_t^\delta dt \right]}{1 - \theta^{1-(\beta/\delta)}}. \quad (A10)$$

In the special case $\delta = 1$ we have, from the Lemma 1, that

$$E_0 \left[\int_0^{\tau_\theta^1} e^{-rt} Y_t^1 dt \right] = (1 - \theta^{1-\beta}) \pi. \quad (A11)$$

Employing the fact that τ_θ^1 and $\tau_{\theta^\delta}^\delta$ have the same distribution, we have that τ_θ^δ is distributed the same as $\tau_{\theta^{1/\delta}}^1$. Now if $\theta = 1 + \epsilon$, where $\epsilon \approx 0$ and positive, then for all $t \leq \tau_\theta^\delta$, $Y_t^\delta \approx Y_t^1$, and

$$E_0 \left[\int_0^{\tau_\theta^\delta} e^{-rt} Y_t^\delta dt \right] = (1 + o(\epsilon)) E_0 \left[\int_0^{\tau_{\theta^{1/\delta}}^1} e^{-rt} Y_t^1 dt \right]. \quad (A12)$$

Substituting into the right hand side of this equation using the preceding equation gives

$$\begin{aligned} E_0 \left[\int_0^{\tau_\theta^\delta} e^{-rt} Y_t^\delta dt \right] &= (1 + o(\epsilon)) \left(1 - \left(\theta^{(1/\delta)} \right)^{1-\beta} \right) \pi \\ &= \left(\frac{\beta-1}{\delta} \right) \epsilon \pi + o(\epsilon^2). \end{aligned} \quad (A13)$$

Substituting into Equation (A10) and taking the limit as ϵ goes to zero yields

$$\pi_\delta = \lim_{\epsilon \rightarrow 0} \frac{\left(\frac{\beta-1}{\delta} \right) \epsilon \pi + o(\epsilon^2)}{\left(\frac{\beta}{\delta} - 1 \right) \epsilon + o(\epsilon^2)} = \left(\frac{\beta-1}{\beta-\delta} \right) \pi. \quad (A14)$$

The parameter restriction assures that $\pi_\delta < \infty$. ■

Proof of the proposition: Rewriting the first term in Equation (8), the value of the intermediate cash flows, as the difference between two perpetual cash flows, one beginning immediately and the other beginning on the date of redevelopment, we have

$$E_t \left[\int_t^\tau e^{-r(s-t)} q P_s ds \right] = \Pi q P_t - E_t \left[e^{-r(\tau-t)} \right] \Pi q P_q^* \quad (A15)$$

where $\tau \equiv \min\{s \geq t | P_s = P_q^*\}$ is the first time prices hit the development threshold, and $\Pi \equiv E_t \left[\int_t^\infty e^{-r(s-t)} \left(\frac{P_s}{P_t} \right) ds \right]$ is the dollar price of a unit cash flow, at today's price, derived from a flow of the firm's output.

The dollar price of a unit cash flow follows immediately from Lemma 5. The assumed price process follows partially reflected geometric Brownian motion with $\delta = \beta/\eta$, so

$$\Pi = \pi_{\beta/\eta} = \frac{\beta-1}{\beta-\frac{\beta}{\eta}} \pi = \frac{(\beta-1)\eta}{\beta(\eta-1)} \pi. \quad (A16)$$

The state price for the first passage of prices to the development threshold, $E_t \left[e^{-r(\tau-t)} \right]$, can be calculated by using the standard results on the underlying process. We can determine the demand that will result in prices reaching the development price threshold P_q^* quite easily, because we know the evolution of the price of firms' output as a function of demand. Inspection of the price--demand relation derived in Proposition 2, using $X_t = \bar{X}_t$ and

$X_\tau = \bar{X}_\tau$, yields

$$P_\tau = \left(\frac{X_\tau}{X_t} \right)^{\beta/\eta} P_t. \quad (\text{A17})$$

This implies, using $P_\tau = P_q^*$, that $(X_\tau/X_t) = (P_q^*/P_t)^{\eta/\beta}$. Then using the fact that the demand process is geometric Brownian we have

$$E_t [e^{-r(\tau-t)}] = \left(\frac{X_t}{X_\tau} \right)^\beta = \left(\frac{P_t}{P_q^*} \right)^\eta. \quad (\text{A18})$$

Substituting Equation (A15) for the first term in Equation (8) and using the above values for $E_t [e^{-r(\tau-t)}]$ and Π yields the proposition. ■

Proof of Proposition 4.

lemma 3 . Suppose that $d \ln P_t$ is a homogeneous function of degree zero in P_t and $\bar{P}_t$. Then the value-at-max function has the following scaling property:

$$W(q, P) = q^\gamma W(1, q^{1-\gamma} P). \quad (\text{A19})$$

Proof of lemma: Suppose a firm with existing capacity q , when $P_t = P$ and $\bar{P}_t = \bar{P}$, optimally develops to capacity q_i at times $\tau_i \equiv \min\{s > t | P_s = T_i^*\}$, where $\{T_i^*\}$ is a sequence of price thresholds and $i = 1, 2, \dots$

Now consider the firm with existing capacity 1 when $P_t = q^{1-\gamma} P$ and $\bar{P}_t = q^{1-\gamma} \bar{P}$. Suppose this firm chooses the investment policy of developing to $\hat{q}_i \equiv q_i/q$ at times $\hat{\tau}_i \equiv \min\{s > t | P_s = q^{1-\gamma} P_{q_i}^*\}$. Then

$$\begin{aligned} V(1, q^{1-\gamma} P_t, q^{1-\gamma} \bar{P}_t) &\geq E_t \left[\int_t^\infty e^{-r(s-t)} \left(\frac{q_s}{q} \right) (q^{1-\gamma} P_s) ds - \sum_{i=1}^\infty e^{-r(\hat{\tau}_i-t)} \left(\frac{q_s}{q} \right)^\gamma \right] \\ &= q^{-\gamma} V(q, P_t, \bar{P}_t) \end{aligned}$$

where $q_s \equiv q_i$ if $s \in [\hat{\tau}_i, \hat{\tau}_{i+1})$, the inequality reflects the fact that the policy may be suboptimal, and the equality holds because $\hat{\tau}_i$ and τ_i have the same distribution.

A symmetric argument, assuming a firm with capacity q replicates the optimal investment strategy of the firm with unit capacity, scaled appropriately, yields the reverse inequality, $V(q, P_t, \bar{P}_t) \geq q^\gamma V(1, q^{1-\gamma} P_t, q^{1-\gamma} \bar{P}_t)$. Together the two inequalities imply

$$V(q, P, \bar{P}_t) = q^\gamma W(1, q^{1-\gamma} P, q^{1-\gamma} \bar{P}_t), \quad (\text{A20})$$

and the value-at-max function inherits the scaling condition directly from equation (A20). ■

Proof of the proposition: Lemma 3 implies a relation between the investment behavior of an arbitrary firm and that of the firm with unit capacity, and consequently between the investment behaviors of two arbitrary firms.¹¹ In particular, it implies a multiplicative investment rule. If the firm with unit capacity optimally redevelops to capacity $q_1^* = \kappa \times (\text{unit capacity})$ then a firm with capacity q optimally redevelops to capacity $q_q^* = kq$, for some k . It is sufficient, consequently, to restrict attention to the strategy of a firm with unit capacity.

Substituting the functional form given in Equation (10), $W(q, P) = \Pi q P + a_q P^\eta$, into the value matching and smooth pasting conditions, Equations (11) and (12), and assuming

¹¹ The investment behavior relationship implied by Lemma 3 can be used to motivate the Strategy Hypothesis.

the firm develops to k times existing density, yields

$$\begin{aligned}\Pi(k-1)P_1^* &= (a_1 - a_k) P_1^{*\eta} + k^\gamma \\ \Pi(k-1) &= \eta (a_1 - a_k) P_1^{*(\eta-1)}.\end{aligned}$$

Solving these immediately yields

$$(a_1 - a_k) = \frac{k^\gamma}{(\eta-1)P_1^{*\eta}} \quad (\text{A21})$$

$$\begin{aligned}P_1^* &= \frac{\eta}{(\eta-1)\Pi} \frac{k^\gamma}{(k-1)} \\ &= \frac{\beta}{(\beta-1)\pi} \frac{k^\gamma}{(k-1)}.\end{aligned} \quad (\text{A22})$$

We need one additional constraint to pin down the three variables. The standard constraint, common in the literature, is that $a_k = 0$. This constraint is generated by assuming a project may be developed one time only, so is unavailable here.

An appropriate constraint is, however, provided by the scaling condition on the value function, Lemma 3. Because development rights are retained at exercise, the scaling condition holds *across development boundaries*. Using this condition we have

$$k^\gamma W^1(1, k^{1-\gamma} P^*) = W^2(k, P^*). \quad (\text{A23})$$

Using the functional form for W^1 and W^2 given by Equation (10), we then have

$$k^\gamma \left[\Pi(k^{1-\gamma} P^*) + a_1 (k^{1-\gamma} P^*)^\eta \right] = \Pi k P^* + a_k P^{*\eta}, \quad (\text{A24})$$

which, solving for a_k , yields

$$a_k = k^{\gamma+\eta-\gamma\eta} a_1. \quad (\text{A25})$$

In conjunction with the equations for $(a_1 - a_k)$ and P_1^* this determines the value of a project in terms of the choice variable, k . In particular, we have that

$$\begin{aligned}a_1 &= \frac{k^\gamma}{(1-k^{\gamma+\eta-\gamma\eta})(\eta-1)P_1^{*\eta}} \\ &= \frac{1}{\eta-1} \left(\frac{(\beta-1)\pi}{\beta} \right)^\eta \frac{(k-1)^\eta}{k^\eta (k^{\gamma\eta-\gamma-\eta-1})}.\end{aligned} \quad (\text{A26})$$

The optimal strategy is the one that maximizes firm value, so the redevelopment capacity must maximize a_1 , so

$$\kappa = \operatorname{argmax}_k \left\{ \frac{(k-1)^\eta}{k^\eta (k^{\gamma\eta-\gamma-\eta-1})} \right\}. \quad (\text{A27})$$

Finally, let $F(x) = \frac{(x-1)^\eta}{x^\eta (x^{\gamma\eta-\gamma-\eta-1})}$. Now κ is greater than one and maximizes $F(x)$, so $F'(\kappa) = 0$. Simple algebra shows that $F'(\kappa) = 0$ if and only if

$$(\gamma\eta - \gamma - \eta)x^{\gamma\eta-\gamma-\eta+1} - (\gamma\eta - \gamma)x^{\gamma\eta-\gamma-\eta} + \eta = 0. \quad (\text{A28})$$

■

Proof of Proposition 5. *Proof of the proposition:* If $d \ln P_t$ is a homogeneous, degree-zero function of P_t and $\bar{P}_t$ then, by Lemma 3, $V(q, P_t, \bar{P}_t) = q^\gamma V(1, q^{1-\gamma} P_t, q^{1-\gamma} \bar{P}_t)$. This scaling condition on the value function implies the optimal investment strategy is a multiplicative strategy of building to k times existing capacity when prices reach $k^{\gamma-1}$ times the price at

which the existing capacity was itself developed, for some k . This sort of rule, for any k , implies the price process of Proposition 2.

Proof of Proposition 6. Homogeneous firms will not choose to undertake the same investment at the same time, because if all firms were to follow such a strategy an individual firm would find it more profitable to deviate, either by developing sooner than other firms or by delaying investment.

Suppose firms do follow a symmetric strategy. That is, suppose firms are homogeneous, possessing the same initial capacities q , and follow the same development strategy. That is, at some point all firms will develop the same new capacity at the same time. Until this time aggregate capacity is fixed, so the price of firms' output evolves like the multiplicative demand shock, as a geometric Brownian motion. The investment problem faced by firms is therefore Markovian up until the time of development, so the investment strategy may be characterized as a trigger strategy on the price of firms' output. That is, there exist some P^* such that the first time P_t , the price of firms' output, reaches P^* all firms add new capacity, developing to the new capacity q^* .

Because the cost of developing new capacity includes the opportunity cost of abandoning old capacity, individual firms do not undertake incremental investment, that is, q^* is bounded away from q . Aggregate capacity would also be lumpy (discontinuous) in a symmetric equilibrium, so the price of firms' output immediately after development is strictly less than the trigger price, that is, $P_{\tau+} < P^*$, where $\tau \equiv \min\{t > 0 | P_t = P^*\}$.

Because the opportunity cost for firms is higher after they have developed to the new, larger capacity q^* , no further development will happen until prices reach some price which is strictly higher than P^* . Let $V_{s,t}$ denote the value of the firm at time s when development to capacity q^* will be undertaken at time t . Then the time- τ value of the firm if it develops to capacity q^* at τ is

$$V_{\tau,\tau} = q^* P_{\tau+} \pi - C(q^*) + \left(\frac{P_{\tau+}}{P^*} \right)^\beta (V_{\tau^*} - q^* P^* \pi), \quad (\text{A29})$$

where $\pi = \frac{1}{r-\mu}$ is the standard geometric Brownian annuity factor, $\tau^* \equiv \min\{t > \tau | P_t = P^*\}$ is the second passage time for the price process to P^* , and V_{τ^*} is the value of having capacity q^* at time τ^* , and we have used $E_\tau \left[e^{-r(\tau^*-\tau)} \right] = \left(\frac{P_{\tau+}}{P^*} \right)^\beta$.

The time- τ value of the firm if it delays development to capacity q^* until τ^* is

$$V_{\tau,\tau^*} = q P_{\tau+} \pi + \left(\frac{P_{\tau+}}{P^*} \right)^\beta (V_{\tau^*} - C(q^*) - q P^* \pi). \quad (\text{A30})$$

The difference, $V_{\tau,\tau^*} - V_{\tau,\tau}$, is the time- τ value of delaying the investment, and is

$$C(q^*) \left(1 - \left(\frac{P_{\tau+}}{P^*} \right)^\beta \right) - (q^* - q) P_{\tau+} \pi \left(1 - E_\tau \left(\frac{P_{\tau+}}{P^*} \right)^{\beta-1} \right). \quad (\text{A31})$$

That is, delaying investment has the advantage of deferring the investment cost, but the disadvantage of loosing some intermediate cash flows. Whether the net benefit is positive or negative depends on the relative magnitudes of the two effects.

An identical argument shows that the value of investing early, to capacity q^* at time $\tau_* \equiv \min\{t > 0 | P_t = P_{\tau+}\}$, is $V_{\tau_*,\tau_*} - V_{\tau_*,\tau} = -(V_{\tau_*,\tau_*} - V_{\tau,\tau})$. Consequently, unless an investor is indifferent between developing to capacity q^* at times τ_* , τ , and τ^* , she has a strict preference for deviating from the symmetric strategy, and benefits from developing either early or late.

If she is indifferent between developing at τ_* and τ , however, she has a strict preference for developing at $\tau_*^* \equiv \min\{t > 0 | P_t = \sqrt{P^* P_{\tau+}}\}$. The value of delaying from τ_* and τ_*^* is

$$\begin{aligned} C(q^*) \left(1 - \left(\frac{P_{\tau+}}{P^*} \right)^{\beta/2} \right) - (q^* - q) P_{\tau+} \pi \left(1 - \left(\frac{P_{\tau+}}{P^*} \right)^{(\beta-1)/2} \right) \\ = \left(1 + \left(\frac{P_{\tau+}}{P^*} \right)^{\beta/2} \right)^{-1} \left[C(q^*) \left(1 - \left(\frac{P_{\tau+}}{P^*} \right)^{\beta} \right) - (q^* - q) P_{\tau+} \pi \left(1 - \left(\frac{P_{\tau+}}{P^*} \right)^{\beta-1} \right) \right. \\ \left. + \left(\frac{P_{\tau+}}{P^*} \right)^{\beta/2} \left(\sqrt{\frac{P^*}{P_{\tau+}}} - 1 \right) (q^* - q) P_{\tau+} \pi \left(1 - \left(\frac{P_{\tau+}}{P^*} \right)^{(\beta-1)/2} \right) \right] \\ = \left(1 + \left(\frac{P_{\tau+}}{P^*} \right)^{\beta/2} \right)^{-1} \left(\sqrt{\frac{P^*}{P_{\tau+}}} - 1 \right) (q^* - q) P_{\tau+} \pi \left(1 - \left(\frac{P_{\tau+}}{P^*} \right)^{(\beta-1)/2} \right), \end{aligned}$$

where the second equality follows from the assumption that firms are indifferent between developing at τ and τ_* . Now the right hand side is strictly positive, so the firm strictly prefers investing at τ_*^* to investing at τ_* . The firm is indifferent between developing at τ_* and τ , so the preceding argument also implies that the firm strictly prefers investing at τ_*^* to τ . ■

Proof of Proposition 7. The value of an arbitrary firm is given by

$$\begin{aligned} V(q, P_t, \bar{P}_t) &= E_t \left[\int_t^{\tau} \bar{P}_t e^{-r(s-t)} q P_s ds \right] + E_t \left[e^{-r(\tau - t)} W(q, P_q) \right] \\ &= \pi q P_t - \left(\frac{P_t}{P_t} \right)^{\beta} \pi q \bar{P}_t + \left(\frac{P_t}{P_t} \right)^{\beta} W(q, \bar{P}_t). \end{aligned} \quad (A32)$$

The functional form of the value-at-max is given by equation (10), $W(q, P) = \Pi q P + a_q P^{\eta}$. After substituting for Π Propositions 4 and a_q from the proof of Propositions 4 we have

$$W(q, P) = \Pi q P + (a_1 P_1^{*\eta}) q^{\gamma} \left(\frac{q^{1-\gamma} P}{P_1^*} \right)^{\eta}. \quad (A33)$$

Substituting this into the previous equation with $P = \bar{P}_t$, and using $a_1 P_1^{*\eta} = \frac{\kappa^{\gamma}}{(\eta-1)(1-\kappa)^{\gamma+\eta-\gamma\eta}}$, yields the proposition, where the parameter restriction guarantees that firms have finite value.

Proof of Proposition 8.

lemma 4. For any $\phi > 0$ and $\gamma > 1$, if $y > 1$ satisfies $\phi y^{\phi+1} - \left(\phi + \frac{\phi+\gamma}{\gamma-1} \right) y^{\phi} + \left(\frac{\phi+\gamma}{\gamma-1} \right) = 0$ then

$$\frac{y^{\phi}-1}{\phi} > \left(1 + \frac{\phi}{\gamma} \right) \ln y. \quad (A34)$$

Proof of lemma: for all $x > 1$ and any $\phi > 0$

$$\frac{d}{dx} (x^{\phi+1} - (\phi+1)x + \phi) = (\phi+1) (x^{\phi} - 1) > 0$$

$$\text{and } \lim_{x \rightarrow 1^+} (x^{\phi+1} - (\phi+1)x + \phi) = 0,$$

$$\Rightarrow x^{\phi+1} - (\phi+1)x + \phi > 0.$$

So for all $x > 1$ and any $\phi > 0$

$$\begin{aligned} & \frac{d}{dx} \left[x \left(\frac{x^\phi - 1}{\phi} \right) + \left(\frac{x^{\phi+1} - 1}{\phi+1} \right) - (x-1) - (\phi+1)x \ln x + \phi \ln x \right] \\ &= \left(\frac{x^\phi - 1}{\phi} - \ln x \right) + \phi \left(\frac{x^{\phi+1} - 1}{\phi+1} - \ln x \right) + x^{-1} \left(x^{\phi+1} - (\phi+1)x + \phi \right) > 0 \\ \text{and } \lim_{x \rightarrow 1^+} & \left(x \left(\frac{x^\phi - 1}{\phi} \right) + \left(\frac{x^{\phi+1} - 1}{\phi+1} \right) - (x-1) - (\phi+1)x \ln x + \phi \ln x \right) = 0, \\ \Rightarrow & x \left(\frac{x^\phi - 1}{\phi} \right) + \left(\frac{x^{\phi+1} - 1}{\phi+1} \right) - (x-1) - (\phi+1)x \ln x + \phi \ln x > 0. \end{aligned}$$

So for all $x > 1$ and any $\phi > 0$

$$\begin{aligned} & \frac{d}{dx} \left[\left(\frac{x^{\phi+1} - 1}{\phi+1} \right) \left(\frac{x^\phi - 1}{\phi} \right) - x^\phi (x-1) \ln x \right] \\ &= x^{\phi-1} \left[x \left(\frac{x^\phi - 1}{\phi} \right) + \left(\frac{x^{\phi+1} - 1}{\phi+1} \right) - (x-1) - (\phi+1)x \ln x + \phi \ln x \right] > 0 \\ \text{and } \lim_{x \rightarrow 1^+} & \left(\left(\frac{x^{\phi+1} - 1}{\phi+1} \right) \left(\frac{x^\phi - 1}{\phi} \right) - x^\phi (x-1) \ln x \right) = 0, \\ \Rightarrow & \left(\frac{x^{\phi+1} - 1}{\phi+1} \right) \left(\frac{x^\phi - 1}{\phi} \right) - x^\phi (x-1) \ln x > 0, \end{aligned}$$

or

$$\frac{x^\phi - 1}{\phi} > \left(\frac{(\phi+1)x^\phi (x-1)}{x^{\phi+1} - 1} \right) \ln x. \quad (\text{A35})$$

This holds for all $x > 1$, so holds for $y > 1$ that satisfies $\phi y^{\phi+1} - \left(\phi + \frac{\phi+y}{y-1} \right) y^\phi + \left(\frac{\phi+y}{y-1} \right) = 0$, which also satisfies

$$(\phi+1)y^\phi(y-1) = \left(1 + \frac{\phi}{y} \right) (y^{\phi+1} - 1). \quad (\text{A36})$$

Substitution into Equation (A35) yields the lemma. ■

Proof of the proposition: Let $f(x) = (\gamma\eta - \gamma - \eta)x^{\gamma\eta - \gamma - \eta + 1} - (\gamma\eta - \gamma)x^{\gamma\eta - \gamma - \eta} + \eta$. Because $f'(x) = 0$ if and only if $x = 0$ or $x = \frac{\gamma\eta - \gamma}{\gamma\eta - \gamma - \eta + 1}$, which is strictly greater than 1, $f(x)$ can have at most two positive solutions. Since $x = 1$ is a solution, it can have at most one other. Since $f\left(\frac{\gamma\eta - \gamma}{\gamma\eta - \gamma - \eta + 1}\right) = -\left(1 + \frac{\eta - 1}{\gamma\eta - \gamma - \eta + 1}\right)^{\gamma\eta - \gamma - \eta + 1} + \eta < 0$, and $f(x)$ is clearly positive for sufficiently large x , $f(x)$ must have a root greater than $\frac{\gamma\eta - \gamma}{\gamma\eta - \gamma - \eta + 1} = \frac{\gamma}{\gamma - 1}$. That is, $f(x)$ has a unique root greater than 1.

Now the curvature parameter η is strictly increasing in r and strictly decreasing in μ , σ and α (i.e., $\frac{d\eta}{dr} > 0$, $\frac{d\eta}{d\mu} < 0$, $\frac{d\eta}{d\sigma} < 0$, $\frac{d\eta}{d\alpha} < 0$), so it is sufficient to show $\frac{d\kappa}{d\eta} < 0$. Now $\frac{df}{dx} > 0$ for all $x > \frac{\gamma}{\gamma - 1}$, so to show $\frac{d\kappa}{d\eta} < 0$ it is sufficient, because $f(\kappa) = 0$ and $\kappa > \frac{\gamma}{\gamma - 1}$, to show that $\frac{\partial f}{\partial \eta} \Big|_{\kappa} > 0$. Now

$$\begin{aligned} & \frac{\partial}{\partial \eta} \left[(\gamma\eta - \gamma - \eta)x^{\gamma\eta - \gamma - \eta + 1} - (\gamma\eta - \gamma)x^{\gamma\eta - \gamma - \eta} + \eta \right]_{x=q_1^\kappa} \\ &= \{(\gamma - 1) + (\gamma\eta - \gamma - \eta)[(\gamma - 1) \ln \kappa]\} q_1^{\kappa\gamma\eta - \gamma - \eta + 1} \\ &\quad - \{\gamma + (\gamma\eta - \gamma)[(\gamma - 1) \ln \kappa]\} q_1^{\kappa\gamma\eta - \gamma - \eta} + 1 \\ &= \frac{\gamma}{\gamma\eta - \gamma - \eta} \left(q_1^{\kappa\gamma\eta - \gamma - \eta - 1} - \eta(\gamma - 1) \ln \kappa \right), \end{aligned}$$

so $\frac{d\kappa}{d\eta} < 0$ if and only if

$$\frac{q_1^{\kappa\gamma\eta-\gamma-\eta-1}}{\gamma\eta-\gamma-\eta} > \frac{\eta(\gamma-1)}{\gamma} \ln \kappa,$$

which holds by lemma 4 with $\phi \equiv \gamma\eta - \gamma - \eta$.

To get the lower bound on the development multiple, note that the defining equation for κ may be rewritten as

$$\left(\frac{\kappa^{\gamma\eta-\gamma-\eta}}{\kappa^{\gamma\eta-\gamma-\eta-1}} \right) (\kappa - 1) = \frac{\eta}{\gamma\eta-\gamma-\eta},$$

which in the limit as η gets large says $\kappa - 1 = \frac{1}{\gamma-1}$, that is,

$$\lim_{\eta \rightarrow \infty} \kappa = \frac{\gamma}{\gamma-1}. \quad (\text{A37})$$

To get the upper bound, note that the defining equation may also be rewritten as

$$\kappa^{\gamma\eta-\gamma-\eta} (\kappa - 1) = \left(\frac{\kappa^{\gamma\eta-\gamma-\eta-1}}{\gamma\eta-\gamma-\eta} \right) \eta,$$

which in the limit says

$$\lim_{\eta \rightarrow \frac{\gamma}{\gamma-1}} \frac{\kappa - 1}{\ln \kappa} = \frac{\gamma}{\gamma - 1}. \quad (\text{A38})$$

■

Proof of Proposition 9.

lemma 5. Suppose X_t follows geometric Brownian motion with drift μ and volatility σ . Then $E_0 \left[\left(\frac{X_t}{X_0} \right) \left(\frac{\bar{X}_0}{\bar{X}_t} \right)^\Omega \right]$ is equal to

$$\begin{aligned} e^{\mu t} \left(N \left(\frac{\ln(\bar{X}_0/X_0) - \left(\mu + \frac{\sigma^2}{2} \right) t}{\sigma \sqrt{t}} \right) + \left(\frac{\Omega}{1 + \frac{2\mu}{\sigma^2} - \Omega} \right) N \left(\frac{-\ln(\bar{X}_0/X_0) - \left(\mu + \frac{\sigma^2}{2} \right) t}{\sigma \sqrt{t}} \right) \right) \left(\frac{\bar{X}_0}{X_0} \right)^{\left(1 + \frac{2\mu}{\sigma^2} \right)} \\ + \left(1 - \frac{\Omega}{1 + \frac{2\mu}{\sigma^2} - \Omega} \right) e^{-\Omega \left(\mu + (1-\Omega) \frac{\sigma^2}{2} \right) t} N \left(\frac{-\ln(\bar{X}_0/X_0) + \left(\mu + (1-2\Omega) \frac{\sigma^2}{2} \right) t}{\sigma \sqrt{t}} \right) \left(\frac{\bar{X}_0}{X_0} \right)^\Omega \Bigg) \quad (\text{A39}) \end{aligned}$$

Proof of lemma: X_t/X_0 is distributed $e^{\mu t + \sigma \sqrt{t} \chi}$, where χ is the standard normal. In conjunction with the joint density for the value and the maximum of a standard Brownian up to time t , $v(b, m) = \sqrt{2/\pi} (2m - b)/t \sqrt{t} e^{-(2m-b)^2/2t}$, and a change of measure, we can write $E_0 \left[(X_t/X_0) (\bar{X}_0/\bar{X}_t)^\Omega \right]$ in integral form as

$$\int_{m=0}^{\infty} \int_{b=-\infty}^m e^{\sigma b} \left(\frac{\bar{X}_0}{\max[X_0, X_0 e^{\sigma m}]} \right)^\Omega e^{\left(\frac{\mu}{\sigma} - \frac{\sigma}{2} \right) b - \left(\frac{\mu}{\sigma} - \frac{\sigma}{2} \right)^2 \frac{t}{2}} \sqrt{\frac{2}{\pi}} \frac{2m-b}{t \sqrt{t}} e^{\frac{-(2m-b)^2}{2t}} db dm. \quad (\text{A40})$$

Rearranging the previous equation yields

$$e^{-\left(\frac{\mu}{\sigma} - \frac{\sigma}{2} \right)^2 \frac{t}{2}} \int_{m=0}^{\infty} \min \left[1, \left(\frac{\bar{X}_0}{X_0} e^{-\sigma m} \right)^\Omega \right] \int_{b=-\infty}^m e^{\left(\frac{\mu}{\sigma} + \frac{\sigma}{2} \right) b} \sqrt{\frac{2}{\pi}} \frac{2m-b}{t \sqrt{t}} e^{\frac{-(2m-b)^2}{2t}} db dm. \quad (\text{A41})$$

The interior integral, after completing the square in the exponential and simplifying, is

$$e^{\left(\frac{\mu}{\sigma} + \frac{\sigma}{2}\right)^2 \frac{t}{2}} e^{2m\left(\frac{\mu}{\sigma} + \frac{\sigma}{2}\right)} \int_{b=-\infty}^m \left(\sqrt{\frac{2}{\pi t}} e^{\frac{2m-b}{t\sqrt{t}}} e^{\frac{-(b-(2m+(\frac{\mu}{\sigma} + \frac{\sigma}{2})t))^2}{2t}} \right) db, \quad (\text{A42})$$

which may be evaluated using $\frac{2m-b}{t\sqrt{t}} = -\frac{1}{\sqrt{t}} \left(\frac{b-(2m+(\frac{\mu}{\sigma} + \frac{\sigma}{2})t)}{t} \right) - \frac{\mu}{\sigma} + \frac{\sigma}{2}$, which yields

$$e^{\left(\frac{\mu}{\sigma} + \frac{\sigma}{2}\right)^2 \frac{t}{2} + \left(\frac{\mu}{\sigma} + \frac{\sigma}{2}\right)2m} \left(\sqrt{\frac{2}{\pi t}} e^{\frac{-(m+(\frac{\mu}{\sigma} + \frac{\sigma}{2})t)^2}{2t}} - 2\left(\frac{\mu}{\sigma} + \frac{\sigma}{2}\right) N\left(\frac{-m-(\frac{\mu}{\sigma} + \frac{\sigma}{2})t}{\sqrt{t}}\right) \right). \quad (\text{A43})$$

Substituting into Equation (A41) gives

$$e^{\mu t} \left(\int_0^{\frac{1}{\sigma} \ln(\frac{\bar{X}_0}{X_0})} e^{2\left(\frac{\mu}{\sigma} + \frac{\sigma}{2}\right)m} \left(\sqrt{\frac{2}{\pi t}} e^{\frac{-(m+(\frac{\mu}{\sigma} + \frac{\sigma}{2})t)^2}{2t}} - 2\left(\frac{\mu}{\sigma} + \frac{\sigma}{2}\right) N\left(\frac{-m-(\frac{\mu}{\sigma} + \frac{\sigma}{2})t}{\sqrt{t}}\right) \right) dm \right. \\ \left. + \left(\frac{\bar{X}_0}{X_0}\right)^\Omega \int_{\frac{1}{\sigma} \ln(\frac{\bar{X}_0}{X_0})}^\infty e^{(2(\frac{\mu}{\sigma} + \frac{\sigma}{2}) - \sigma\Omega)m} \left(\sqrt{\frac{2}{\pi t}} e^{\frac{-(m+(\frac{\mu}{\sigma} + \frac{\sigma}{2})t)^2}{2t}} - 2\left(\frac{\mu}{\sigma} + \frac{\sigma}{2}\right) N\left(\frac{-m-(\frac{\mu}{\sigma} + \frac{\sigma}{2})t}{\sqrt{t}}\right) \right) dm \right). \quad (\text{A44})$$

Evaluating the integral over the region $0 < m < \frac{1}{\sigma} \ln(\bar{X}_0/X_0)$, the first term directly and the second term by integration by parts, we get

$$N\left(\frac{\ln(\bar{X}_0/X_0) - (\mu + \frac{\sigma^2}{2})t}{\sigma\sqrt{t}}\right) - N\left(\frac{-\ln(\bar{X}_0/X_0) - (\mu + \frac{\sigma^2}{2})t}{\sigma\sqrt{t}}\right) \left(\frac{\bar{X}_0}{X_0}\right)^{\left(1 + \frac{2\mu}{\sigma^2}\right)}. \quad (\text{A45})$$

Evaluating the integral over the region $0 < m < \frac{1}{\sigma} \ln(\bar{X}_0/X_0)$, the first term directly and the second term by integration by parts, we get

$$N\left(\frac{\ln(\bar{X}_0/X_0) - (\mu + \frac{\sigma^2}{2})t}{\sigma\sqrt{t}}\right) - N\left(\frac{-\ln(\bar{X}_0/X_0) - (\mu + \frac{\sigma^2}{2})t}{\sigma\sqrt{t}}\right) \left(\frac{\bar{X}_0}{X_0}\right)^{\left(1 + \frac{2\mu}{\sigma^2}\right)}. \quad (\text{A46})$$

Evaluating the integral over the region $m > \frac{1}{\sigma} \ln(\bar{X}_0/X_0)$, the first term by completing the square in the exponent and the second term again by parts, we get

$$\left(1 + \frac{\Omega}{1 + \frac{2\mu}{\sigma^2} - \Omega}\right) N\left(\frac{-\ln(\bar{X}_0/X_0) - (\mu + \frac{\sigma^2}{2})t}{\sigma\sqrt{t}}\right) \left(\frac{\bar{X}_0}{X_0}\right)^{\left(1 + \frac{2\mu}{\sigma^2} - \Omega\right)} \\ + \left(1 - \frac{\Omega}{1 + \frac{2\mu}{\sigma^2} - \Omega}\right) e^{-\Omega\left(\mu + (1-\Omega)\frac{\sigma^2}{2}\right)t} N\left(\frac{-\ln(\bar{X}_0/X_0) + (\mu + (1-2\Omega)\frac{\sigma^2}{2})t}{\sigma\sqrt{t}}\right). \quad (\text{A47})$$

Substituting Equations (A45) and (A47) into Equation (A44) yields the lemma.

Proof of the proposition: The forward prices is the expected future price so

$$F_{0,t} = E_0 \left[P_0 \left(\frac{X_t}{X_0} \right) \left(\frac{\bar{X}_0}{\bar{X}_t} \right)^{\frac{1}{1+\alpha(\gamma-1)}} \right], \quad (\text{A48})$$

and the proposition then follows directly from lemma 2, with $\Omega = \frac{1}{1+\alpha(\gamma-1)}$.

Proof of Proposition 10. Heuristically, away from maxima in the price of firms' output log-changes in the output price are normal at sufficiently short time intervals. As a result, returns to a firm, which has a price that is convex in the price of its output, is also convex. That is, if firm value is given by $V(P)$, where P is the price of the firms' output, then because V is convex, are firms' returns, and these are given by $\ln V(P) = \ln V(e^{\ln P})$. Because $\ln P$ is normal, large upward moves in $\ln V(e^{\ln P})$ are more likely than large downward moves.

More formally, we will show the result by perturbation analysis. We will consider the third center moment to returns as the time interval gets very small and show that this is, up to dominating second order terms in dt , positive. If $\bar{P}_t > P_t$ then the probability that prices of firms' output returns to the maximum on the interval from t to $t + dt$ is given by

$$\begin{aligned} P(\bar{P}_{t+dt} > \bar{P}_t) &= P(\bar{X}_{t+dt} > \bar{X}_t) \\ &= 2P(X_{t+dt} > \bar{X}_t) \\ &= o\left(N\left(\frac{X_t - \bar{X}_t}{\sigma\sqrt{dt}}\right)\right), \end{aligned} \quad (\text{A49})$$

and can be ignored in the second order expansion for sufficiently small dt . Because the instantaneous return is essentially normal we will without loss of generality ignore the mean when computing higher order centered moments. That is, for convenience we will assume zero drift in demand (it is of course not zero, but the drift does not effect the higher order centered moments).

By inspection of Equation (13), the decomposition of the value of a firm into cash flows until the price of firms' output returns to the historical maximum and the value of all cash flows after that, we have

$$V(P) = c_1 P + c_2 P^\beta. \quad (\text{A50})$$

where c_1 , c_2 and β are all strictly positive.

Now firm returns are given by $d \ln V(P_t)$, and $d \ln V(P_t) = d \ln(kV(P_t))$ for any constant k . So let us consider $f(x) = \frac{1}{1+c}x + \frac{c}{1+c}x^\beta$, for some c that makes f proportional to V . To second order expansion around 1 we have

$$f(x) \approx 1 + \left(\frac{1+\beta c}{1+c}\right)x + \left(\frac{\beta(\beta-1)c}{1+c}\right)\frac{(x-1)^2}{2}. \quad (\text{A51})$$

Using $e^y \approx 1 + y + \frac{y^2}{2}$ we then have, to second order, that

$$\begin{aligned} f(e^x) &\approx 1 + \left(\frac{1+\beta c}{1+c}\right)\left(x + \frac{x^2}{2}\right) + \left(\frac{\beta(\beta-1)c}{1+c}\right)\frac{x^2}{2} \\ &= 1 + \left(\frac{1+\beta c}{1+c}\right)x + \left(\frac{1+\beta c + \beta(\beta-1)c}{1+c}\right)\frac{x^2}{2}. \end{aligned} \quad (\text{A52})$$

Finally, using $\ln(1+y) \approx y - \frac{y^2}{2}$ we have

$$\begin{aligned} \ln f(e^x) &\approx \left[\left(\frac{1+\beta c}{1+c}\right)x + \left(\frac{1+\beta c + \beta(\beta-1)c}{1+c}\right)\frac{x^2}{2}\right] - \left(\frac{1+\beta c}{1+c}\right)^2 \frac{x^2}{2} \\ &= \left(\frac{1+\beta c}{1+c}\right)x + \left(\frac{\beta(\beta-1)^2 c}{(1+c)^2}\right)\frac{x^2}{2}. \end{aligned} \quad (\text{A53})$$

That is, for small x we have, ignoring terms higher than second order, that $\ln f(e^x) = ax + bx^2$ where a and b are both strictly positive. Using $d \ln P_t$ for x , and the fact that $d \ln P_t$ is normally distributed with variance $\sigma^2 dt$, we have that the third centered moment to stock returns is given by

$$\begin{aligned} \mathbf{E} \left[(\ln f(e^x) - \mathbf{E}[\ln f(e^x)])^3 \right] &= \mathbf{E} \left[\left(ax + bx^2 - \mathbf{E}[ax + bx^2] \right)^3 \right] \\ &= \mathbf{E} \left[\left(ax + b(x^2 - \sigma^2 dt) \right)^3 \right] \\ &= 8b^3 \sigma^6 dt^3 + 18a^2 b dt^2. \end{aligned} \quad (\text{A54})$$

In the limit as dt gets very small, stock returns' third centered moment goes to $18a^2 b dt^2$. This is strictly positive because a and b are both strictly positive. We have, therefore, that if $P_t < \bar{P}_t$ then at sufficiently short horizons stock prices are positively skewed.

When $P_t = \bar{P}_t$ the negative skewness in stock returns is inherited directly from the negative skewness in the price of their output. The price of firms' output is not $It\theta$, and is negatively skewed even in the infinitesimal when $P_t = \bar{P}_t$. Because to order dt stock prices are linear in the price of firms's output, negative skewness in the returns to the price of firms' output at short horizons (in the order dt term) translates directly to negative skewness in stock returns at short horizons. ■

Proof of Proposition 11. For the redevelopment density, we need to compare the κ from the optimal strategy equation,

$$\kappa = \operatorname{argmax}_q \left\{ \frac{(q-1)^\eta}{q^\eta (q^\gamma (\eta-1) - \eta - 1)} \right\},$$

to the optimal strategy of a firm in the partial-equilibrium, geometric Brownian price economy. Using the fact that the economy with geometric Brownian output prices is just the limit of the equilibrium economy as the price elasticity gets very large, and letting overscore tildes denote parameters in this economy, we have

$$\tilde{\kappa} = \operatorname{argmax}_q \left\{ \frac{(q-1)^{\tilde{\beta}}}{q^{\tilde{\beta}} (q^{\tilde{\gamma}} (\tilde{\beta}-1) - \tilde{\beta} - 1)} \right\}. \quad (\text{A55})$$

The long-run average price growth, from section (3.2), is $\tilde{\mu} = \frac{\alpha(\gamma-1)}{1+\alpha(\gamma-1)} \left(\mu - \frac{1}{1+\alpha(\gamma-1)} \frac{\sigma^2}{2} \right)$. The long-run average price variance is implied directly from the equation for the equilibrium price process, and the fact that a Brownian process with positive drift stays "close" to its maximum, $\tilde{\sigma}^2 = \left(\frac{\alpha(\gamma-1)}{1+\alpha(\gamma-1)} \right)^2 \sigma^2$. Using these in the formula for these in

$\tilde{\beta} = \sqrt{\left(\frac{\tilde{\mu}}{\tilde{\sigma}^2} - \frac{1}{2} \right)^2 + \frac{2r}{\tilde{\sigma}^2}} - \left(\frac{\tilde{\mu}}{\tilde{\sigma}^2} - \frac{1}{2} \right)$ yields $\tilde{\beta} = \eta$, and then that $\tilde{\kappa} = \kappa$ follows directly.

For the timing of the redevelopment, we will again use the fact that the partial-equilibrium economy looks like the limit of the equilibrium economy as the price elasticity gets very large. Using $\tilde{\beta} = \eta$ and $\tilde{\kappa} = \kappa$ we have that

$$P_1^* = \frac{\beta(\eta-1)}{(\beta-1)\eta} \tilde{P}_1^* > \tilde{P}_1^*. \quad (\text{A56})$$

■

Proof of Proposition 12. The ratio of option value to intrinsic value at any price maxima is given by

$$\begin{aligned} \frac{W(q, P) - \Pi q P}{\Pi q P} &= \frac{\frac{\kappa^\gamma q^\gamma}{(\eta-1)(1-\kappa^{\gamma+\eta-\gamma\eta})} \left(\frac{P}{P_q^*}\right)^\eta}{\Pi q P} \\ &= \frac{\frac{\kappa^\gamma}{(\eta-1)\Pi P_1^* (1-\kappa^{\gamma+\eta-\gamma\eta})} \left(\frac{P}{P_q^*}\right)^{\eta-1}}{\Pi q P} \\ &= \frac{\frac{\kappa-1}{\eta(1-\kappa^{\gamma+\eta-\gamma\eta})} \left(\frac{P}{P_q^*}\right)^{\eta-1}}{\Pi q P}. \end{aligned} \quad (\text{A57})$$

The proposition then follows from the fact that $\tilde{\beta} = \eta$ (please see the proof of Proposition 11). ■

Proof of Proposition 13. The monopolist will, for the same reasons as competitive firms, follow a multiplicative investment strategy. The time stationarity in the investment decision (the price process is Markovian on any time interval that does not include new price maxima), and the scaling property of the value function (Lemma 3), imply that the monopolist will follow a rule of developing to a fixed multiple, which we will denote $\bar{\kappa}$, times existing capacity when price reach $\bar{\kappa}^{\gamma-1}$ times the price level at which the current capacity was developed. The strategy is then fully specified if we know $\bar{\kappa}$ and M_1^* , because under the multiplicative strategy

$$\begin{aligned} \bar{q}_q^* &= \bar{\kappa} q \\ M_q^* &= q^{\gamma-1} M_1^* \end{aligned}$$

Calculating the optimal development multiple $\bar{\kappa}$ for the monopolist is easier than it was for competitive agents. Whereas competitive agents jointly determine the optimal combination of development density and timing, the monopolist makes the density decision independent of the other margin. The monopolist fully internalizes the cost that the development of one project imposes on the revenues of the other projects, so will provide any new capacity as efficiently (cheaply) as possible.

To determine the most efficient development multiple, assume that the monopolist follows a strategy of developing to an arbitrary multiple f of current capacity, and assume, without loss of generality, that $Q = 1$. Assuming the consistent initial cross-sectional distribution of initial capacities, a log-uniform distribution from the minimum initial capacity $q^m = \frac{\ln f}{f-1}$ to the maximum initial capacity $q^M = \frac{f \ln f}{f-1}$, then the marginal cost of adding aggregate capacity is

$$\begin{aligned} \frac{d}{dy} \int_{q^m}^{(1+\epsilon)q^m} c(fy) dv(y) &= \frac{d}{dy} \int_{q^m}^{(1+\epsilon)q^m} (fy)^\gamma \frac{d \ln y}{\ln f} \\ &= \frac{f^\gamma (\ln f)^{\gamma-1}}{(f-1)^\gamma}. \end{aligned}$$

The monopolist's optimal policy minimizes the cost of adding new capacity, that is, $\bar{\kappa}$ minimizes the previous equation, so must satisfy

$$\frac{\bar{\kappa} - 1}{\ln \bar{\kappa}} = \frac{\gamma}{\gamma - 1}. \quad (\text{A58})$$

To calculate M_1^* we will now assume that $q_0^m = 1$ and determine the price at which the monopolist optimally next adds capacity. The discounted value of following the multiplicative

strategy, net of discounted development costs, is given by

$$\pi Q_0 P_0 + \left(\frac{M_1^*}{P_0}\right)^{1-\beta} (\varpi - \pi) Q_0 P_0 - \left(\frac{M_1^*}{P_0}\right)^{-\beta} \Gamma,$$

where $\pi = \frac{1}{r-\mu}$ is the geometric Brownian annuity factor, ϖ is the dollar value of a unit revenue flow under the actual revenue process when the price of the good is at a maximum, and Γ is the expected discounted cost of all future development starting from the time of the first development. Maximizing over M_1^* yields

$$M_1^* = \frac{\beta \Gamma}{(\beta - 1)(\varpi - \pi) Q_0}. \quad (\text{A59})$$

To calculate ϖ , let $R_t \equiv Q_t P_t$ denote the instantaneous gross revenue process. Under the multiplicative strategy, and starting at a price maximum, $\frac{P_t}{P_0} = \left(\frac{\bar{X}_t}{X_0}\right)^{\frac{-1}{1+\alpha(\gamma-1)}} \left(\frac{X_t}{X_0}\right)$ and $\frac{Q_t}{Q_0} = \left(\frac{\bar{P}_t}{P_0}\right)^{\frac{1}{\gamma-1}} = \left(\frac{\bar{X}_t}{X_0}\right)^{\frac{\alpha}{1+\alpha(\gamma-1)}}$, so

$$\varpi \equiv E_0 \left[\int_0^\infty e^{-rt} \frac{R_t}{R_0} dt \right] = E_0 \left[\int_0^\infty e^{-rt} \left(\frac{\bar{X}_t}{X_0}\right)^\Omega \left(\frac{X_t}{X_0}\right) dt \right]$$

where $\Omega = \frac{\alpha-1}{1+\alpha(\gamma-1)}$, and by Lemma 5 we then have

$$\varpi = \left(\frac{\beta-1}{\beta-1-\frac{\alpha-1}{1+\alpha(\gamma-1)}} \right) \pi. \quad (\text{A60})$$

To calculate Γ , assume that $q_0^m = 1$ and that the next positive demand shock elicits a supply response (i.e., that $P_0 = M_1^*$). Then, letting dc_t denote the cost of new supply added in the interval t to $t + dt$, we have

$$\Gamma \equiv E_0 \left[\int_{t=0}^\infty e^{-rt} dc_t \right] = E_0 \left[\int_{y=1}^\infty e^{-r\tau_y} c(\bar{\kappa}y) dv(y) \right]$$

where τ_y is the stopping time for the first passage of P_t to M_y^* , the price at which the monopolist optimally develops her project with existing capacity y , $c(\bar{\kappa}y) = (\bar{\kappa}y)^\gamma$ is the cost of developing a site to capacity $\bar{\kappa}y$, and $dv(y) = \frac{dy}{y \ln \bar{\kappa}}$ is the mass of projects in the interval from y to $y + dy$. Letting X_1^* and X_y^* denote the levels of the multiplicative demand process that corresponds to the first passage of the price process to M_1^* and $M_y^* = y^{\gamma-1} M_1^*$, respectively, we have that

$$E_0 [e^{-r\tau_y}] = \left(\frac{X_y^*}{X_1^*} \right)^{-\beta} = \left(\frac{M_y^*}{M_1^*} \right)^{-\eta}$$

so

$$\begin{aligned} \Gamma &= \int_1^\infty y^{-(\gamma-1)\eta} (\bar{\kappa}y)^\gamma \frac{dy}{y \ln \bar{\kappa}} \\ &= \frac{\bar{\kappa}^\gamma}{(\gamma\eta - \gamma - \eta) \ln \bar{\kappa}}. \end{aligned} \quad (\text{A61})$$

Substituting for ϖ and Γ in Equation (A59) using Equations (A60) and (A61), and using $Q_0 = \frac{\bar{\kappa}-1}{\ln \bar{\kappa}}$, then yields M_1^* .

Proof of Proposition 14. At the time of the most recent development it was optimal to develop the smallest capacity site, so $\bar{P}_t = P_{q_t^m}^* = (q_t^m)^{\gamma-1} P_1^* = \left(\frac{\ln \kappa}{\kappa-1} Q_t\right)^{\gamma-1} P_1^*$. In conjunction with the inverse pricing rule, which implies $\bar{P}_t = \bar{X}_t Q_t^{-1/\alpha}$, this yields

$$Q_t = \left[\frac{\kappa-1}{\ln \kappa} \left(\frac{\bar{X}_t}{P_1^*} \right)^{\frac{1}{\gamma-1}} \right]^{\beta/\eta}. \quad (\text{A62})$$

Taken with the corresponding equation for the monopolist economy this yields an equation for $\phi \equiv \hat{Q}_t / Q_t$,

$$\phi = \left[\left(\frac{\ln \kappa / (\kappa-1)}{\ln \bar{\kappa} / (\bar{\kappa}-1)} \right) \left(\frac{P_1^*}{M_1^*} \right)^{\frac{1}{\gamma-1}} \right]^{\beta/\eta} = \left[\left(\frac{\alpha-1}{\alpha} \right) \left(\frac{\frac{\kappa^\gamma (\ln \kappa)^{\gamma-1}}{(\kappa-1)^\gamma}}{\frac{\bar{\kappa}^\gamma (\ln \bar{\kappa})^{\gamma-1}}{(\bar{\kappa}-1)^\gamma}} \right) \right]^{\frac{\beta}{(\gamma-1)\eta}} \quad (\text{A63})$$

The right hand side of Equation (A63), and hence ϕ , is strictly less than one because $\bar{\kappa}$ maximizes $\frac{f^\gamma (\ln f)^{\gamma-1}}{(f-1)^\gamma}$.

Proof of Proposition 15. Let L_t denote the total expected discounted future deadweight loss at time t . This loss is the expected discounted lost consumer surplus from the monopolist restricting supply, less the expected efficiency gains that arise from the monopolist's efficient development of production. We can write the expected loss in terms of S_t , the instantaneous lost consumer surplus, and Γ , the expected cost of all future development when the least developed site has unit capacity. The expected loss is then given by

$$L_t = \int_t^\infty e^{-r(t'-t)} S_{t'} dt' - E_t[e^{-e(\tau-t)}] \left((q_t^m)^\gamma \Gamma - (\hat{q}_t^m)^\gamma \hat{\Gamma} \right) \quad (\text{A64})$$

where τ is the stopping time for the first passage of the price back to the maximum, and variables with hats represent parameters in the monopolist economy. Integrating then yields

$$L_t = \frac{1-\phi^{1-1/\alpha}}{\alpha-1} \pi Q_t P_t - \left(\frac{X_t}{X_t} \right)^\beta \left(\frac{1-\phi^{1-1/\alpha}}{\alpha-1} (\pi-\varpi) Q_t \bar{P}_t - \left(1 - \left(\frac{\hat{q}_t^m}{q_t^m} \right)^\gamma \hat{\Gamma} \right) (q_t^m)^\gamma \Gamma \right). \quad (\text{A65})$$

The inverse pricing rule implies both $\bar{X}_t = Q_t^{1/\alpha} ((q_t)^\gamma)^{-1} P_1^*$ and $\bar{X}_t = \hat{Q}_t^{1/\alpha} ((\hat{q}_t)^\gamma)^{-1} P_1^*$, so

$$\left(\frac{\hat{q}_t^m}{q_t^m} \right)^\gamma = \left(\phi^{-1/\alpha} \frac{P_1^*}{M_1^*} \right)^{\frac{\gamma}{\gamma-1}}. \quad (\text{A66})$$

This, in conjunction with

$$\hat{\Gamma} = \frac{\bar{\kappa}^\gamma / \ln \bar{\kappa}}{\kappa^\gamma / \ln \kappa} = \frac{\alpha-1}{\alpha} \left(\frac{(\bar{\kappa}-1) / \ln \bar{\kappa}}{(\kappa-1) / \ln \kappa} \right) \frac{M_1^*}{P_1^*}, \quad (\text{A67})$$

implies

$$\left(\frac{q_t^m}{\hat{q}_t^m} \right)^\gamma \hat{\Gamma} = \frac{\alpha-1}{\alpha} \phi^{\frac{\alpha-1}{\alpha}}. \quad (\text{A68})$$

We may then write the general loss process as

$$L_t = \frac{1-\phi^{1-1/\alpha}}{\alpha-1} \left(Q_t P_t \pi - \left(\frac{X_t}{X_t} \right)^\beta (\pi-\varpi) Q_t \bar{P}_t \right) - \left(\frac{X_t}{X_t} \right)^\beta \left(1 - \frac{\alpha-1}{\alpha} \phi^{1-1/\alpha} \right) \left(\frac{\ln \kappa}{\kappa-1} Q_t \right)^\gamma \Gamma. \quad (\text{A69})$$

In the simplest case, when $P_t = \bar{P}_t$, we then have

$$L_t = \frac{1-\phi^{1-1/\alpha}}{\alpha-1} \varpi Q_t P_t - \left(1 - \frac{\alpha-1}{\alpha} \phi^{1-1/\alpha}\right) \left(\frac{\ln \kappa}{\kappa-1} Q_t\right)^\gamma \Gamma. \quad (\text{A70})$$

■

References

- Barberis, N., A. Shleifer, and R. Vishny. 1998. A Model of Investor Sentiment. *Journal of Financial Economics* 49:307–43.
- Berk, J., R. Green, and V. Naik. 2004. Valuation and Return Dynamics of New Ventures. *Review of Financial Studies* 17:1–35.
- Caballero, R. J., and R. S. Pindyck. 1996. Uncertainty, Investment, and Industry Evolution. *International Economic Review* 37:641–62.
- Dixit, A. K. 1999. Irreversible Investment and Competition Under Uncertainty, in B. Kaushik, M. Majumdar, and T. Mitra (eds), *Capital, Investment, and Development*. Cambridge, MA: Basil Blackwell.
- Dixit, A. K., and R. S. Pindyck. 1994. *Investment Under Uncertainty*. Princeton, NJ: Princeton University Press.
- Downing, C., and N. Wallace. 2001. A Real Options Approach to Housing Investment, Working Paper.
- Durrett, R. 1996. *Probability: Theory and Examples* (2nd ed.), Duxbury Press at Wadsworth Publishing Company, Belmont, California.
- Evans, S. N. 2000. Snakes and Spiders: Brownian Motion on R-Trees. *Probability Theory and Related Fields* 117:361–86.
- Fudenberg, D., and J. Tirole. 2000. *Game Theory* (7th ed.), MIT Press, Cambridge, Massachusetts.
- Garlappi, L. 2004. Risk Premia and Preemption in R&D Ventures. *Journal of Financial and Quantitative Analysis* 39:843–72.
- Grenadier, S. R. 1995. Valuing Lease Contracts: A Real Options Approach. *Journal of Financial Economics* 38:297–331.
- Grenadier, S. R. 1996. The Strategic Exercise of Options: Development Cascades and Overbuilding in Real Estate Markets. *Journal of Finance* 51:1653–79.
- Grenadier, S. R. 2002. Option Exercise Games: An Application to the Equilibrium Investment Strategy of Firms. *Review of Financial Studies* 15:691–721.
- Hotelling, H. H. 1929. Stability in Competition. *Economic Journal* 39:41–57.
- Kogan, L. 2001. An Equilibrium Model of Irreversible Investment. *Journal of Financial Economics* 62:201–45.
- Kogan, L. 2004. Asset Prices and Real Investment. *Journal of Financial Economics* 73:411–32.
- Lambrecht, B., and W. Perraudin. 2003. Real Options and Preemption Under Incomplete Information. *Journal of Economic Dynamics and Control* 27:619–43.
- Leahy, J. V. 1993. Investment in Competitive Equilibrium: The Optimality of Myopic Behavior. *The Quarterly Journal of Economics* 108:1105–33.
- Lucas, R. E. 1978. Asset Prices in an Exchange Economy. *Econometrica* 46:1429–45.
- Lucas, R. E., and E. C. Prescott. 1971. Investment Under Uncertainty. *Econometrica* 39:659–81.
- Pindyck, R. S. 1988. Irreversible Investment, Capacity Choice, and the Value of the Firm. *American Economic Review* 79:969–85.
- Richard, S. F., and M. Sundaresan. 1981. A Continuous Time Equilibrium Model of Forward Prices and Future Prices in a Multigood Economy. *Journal of Financial Economics* 9:347–71.

- Routledge, B. R., D. J. Seppi, and C. S. Spatt. 2000. Equilibrium Forward Curves for Commodities. *Journal of Finance* 55:1297–338.
- Rubinstein, M. 1976. The Valuation of Uncertain Streams and the Pricing of Options. *Bell Journal of Economics* 7:407–25.
- Smets, F. R. 1991. Exporting versus FDI: The Effect of Uncertainty, Irreversibilities and Strategic Interactions, Working Paper.
- Spatt, C. S., and F. P. Sterbenz. 1985. Learning, Preemption, and the Degree of Rivalry. *Rand Journal of Economics* 16:84–92.
- Titman, S. 1985. Urban Land Prices Under Uncertainty. *American Economic Review* 75:505–14.
- Watson, G. N. 1944. *A Treatise on the Theory of Bessel Functions* (2nd ed.), Cambridge: Cambridge University Press.
- Williams, J. T. 1991. Real Estate Development as an Option. *Journal of Real Estate Finance and Economics* 4:191–208.
- Williams, J. T. 1993. Equilibrium and Options on Real Assets. *Review of Financial Studies* 6:825–50.
- Williams, J. T. 1997. Redevelopment of Real Assets. *Real Estate Economics* 25:387–407.

Copyright of Review of Financial Studies is the property of Society for Financial Studies and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.