QUASAR CLUSTERING AND THE LIFETIME OF QUASARS

PAUL MARTINI[1] AND DAVID H. WEINBERG

Department of Astronomy, 140 West 18th Avenue, Ohio State University, Columbus, OH 43210; martini@astronomy.ohio-state.edu, dhw@astronomy.ohio-state.edu

ABSTRACT

Although the population of luminous quasars rises and falls over a period of $\sim 10^9$ yr, the typical lifetime of individual quasars is uncertain by several orders of magnitude. We show that quasar clustering measurements can substantially narrow the range of possible lifetimes, with the assumption that luminous quasars reside in the most massive host halos. If quasars are long-lived, then they are rare phenomena that are highly biased with respect to the underlying dark matter, while if they are short-lived they reside in more typical halos that are less strongly clustered. For a given quasar lifetime, we calculate the minimum host halo mass by matching the observed space density of quasars, using the Press-Schechter approximation. We use the results of Mo & White to calculate the clustering of these halos, and hence of the quasars they contain, as a function of quasar lifetime. A lifetime of $t_Q = 4 \times 10^7$ yr, the e-folding timescale of an Eddington luminosity black hole with accretion efficiency $\epsilon = 0.1$, corresponds to a quasar correlation length of $r_0 \approx 10 \ h^{-1}$ Mpc in low-density cosmological models at $z = 2$–3; this value is consistent with current clustering measurements, but these have large uncertainties. High-precision clustering measurements from the Two-Degree Field (2dF) and Sloan quasar surveys will test our key assumption of a tight correlation between quasar luminosity and host halo mass, and if this assumption holds, then they should determine t_Q to a factor of 3 or better. An accurate determination of the quasar lifetime will show whether supermassive black holes acquire most of their mass during high-luminosity accretion, and it will show whether the black holes in the nuclei of typical nearby galaxies were once the central engines of high-luminosity quasars.

Subject headings: dark matter — large-scale structure of universe — quasars: general

1. INTRODUCTION

Mounting evidence for the existence of supermassive black holes in the centers of nearby galaxies (recently reviewed by, e.g., Richstone et al. 1998) supports the long-standing hypothesis that quasars are powered by black hole accretion (e.g., Salpeter 1964; Zeldovich & Novikov 1964; Lynden-Bell 1969). However, one of the most basic properties of quasars, the typical quasar lifetime, t_Q, remains uncertain by orders of magnitude. The physics of gravitational accretion and radiation pressure provides one natural timescale, the e-folding time $t_e = M_{\rm BH}/\dot{M} = 4 \times 10^8 \epsilon l$ yr of a black hole accreting mass with a radiative efficiency $\epsilon = L/\dot{M}c^2$ and shining at a fraction $l = L/L_{\rm E}$ of its Eddington luminosity (Salpeter 1964). However, while $\epsilon \sim 0.1$ and $l \sim 1$ are plausible values for a quasar, it is possible that black holes accrete much of their mass while radiating at much lower efficiency, or at a small fraction of $L_{\rm E}$. The task of determining t_Q must therefore be approached empirically.

The observed evolution of the quasar luminosity function imposes a strong upper limit on t_Q of about 10^9 yr, since the whole quasar population rises and falls over roughly this interval (see, e.g., Osmer 1998). The lifetime of individual quasars could be much shorter than the lifetime of the quasar population, however, and lower limits of $t_Q \sim 10^5$ yr rest on indirect arguments, such as the requirement that quasars maintain their ionizing luminosity long enough to explain the proximity effect in the Lyα forest (e.g., Bajtlik, Duncan, & Ostriker 1988; Bechtold 1994). A typical lifetime of $t_Q \sim 10^9$ yr would imply that quasars are rare phenomena, arising in at most a small fraction of high-redshift galaxies. Conversely, a lifetime as low as $t_Q \sim 10^5$ yr would imply that quasars are quite common, suggesting that a large fraction of present-day galaxies went through a brief quasar phase in their youth.

The comoving space density $\Phi(z)$ of active quasars at redshift z is proportional to $t_Q n_H(z)$, where n_H is the comoving space density of quasar hosts. "Demographic" studies of the local black hole population (e.g., Magorrian et al. 1998; Salucci et al. 1999; van der Marel 1999) have opened up one route to determining the typical quasar lifetime: counting the present-day descendants of the quasar central engines in order to estimate $n_H(z)$ and thus constrain t_Q by matching $\Phi(z)$. Roughly speaking, the ubiquity of black holes in nearby galaxies suggests that quasars are common and that t_Q is likely in the range of 10^6–10^7 (e.g., Richstone et al. 1998; Haehnelt, Natarajan, & Rees 1998; Salucci et al. 1999). However, as Richstone et al. (1998) emphasize, the lifetime estimated in this way depends crucially on the way one links the mass of a present-day black hole to the luminosity of a high-redshift quasar, which in turn depends on assumptions about the growth of black hole masses since the quasar epoch via mergers or low-efficiency accretion.

In this paper we propose an alternative route to the quasar lifetime, using measurements of high-redshift quasar clustering. The underlying idea goes back to the work of Kaiser (1984) and Bardeen et al. (1986): in models of structure formation based on gravitational instability of Gaussian primordial fluctuations, the rare, massive objects are highly biased tracers of the underlying mass distribution, while more common objects are less strongly biased. There-

[1] Current address: Carnegie Observatories, 813 Santa Barbara Street, Pasadena, CA 91101; martini@ociw.edu.

fore, a longer quasar lifetime (t_Q) should imply a more clustered quasar population, provided that luminous quasars reside in massive hosts. The specific calculations that we present in this paper use the Press & Schechter (1974; hereafter PS) approximation for the mass function of dark matter halos and the Mo & White (1996, hereafter MW) and Jing (1998) approximations for the bias of these halos as a function of mass. The path from clustering to quasar lifetime has its own uncertainties; in particular, our predictions for quasar clustering rely on the assumption that the luminosity of a quasar during its active phase is a monotonically increasing function of the mass of its host dark matter halo. However, the assumptions in the clustering approach are at least very different from those in the black hole mass function approach, and they can be tested empirically by detailed studies of quasar clustering as a function of luminosity and redshift.

Our theoretical model of quasar clustering follows a general trend in which the study of quasar activity is embedded in the broader context of galaxy formation and gravitational growth of structure (e.g., Efstathiou & Rees 1988; Turner 1991; Haehnelt & Rees 1993; Katz et al. 1994; Haehnelt et al. 1998; Haiman & Loeb 1998; Monaco, Salucci, & Danese 2000; Kauffmann & Haehnelt 2000). This paper also continues a theme that is prominent in recent work on the clustering of Lyman-break galaxies, namely, that the clustering of high-redshift objects is a good tool for understanding the physics of their formation and evolution (e.g., Adelberger et al. 1998; Katz, Hernquist, & Weinberg 1999; Kolatt et al. 1999; Mo, Mao, & White 1999). Our model of the quasar population is idealized, but by focusing on a simple calculation with clearly defined predictions, we hope to highlight the link between quasar lifetime and clustering strength. After presenting the theoretical results, we draw some inferences from existing estimates of the quasar correlation length. However, our study is motivated mainly by an anticipation of vastly improved measurements of quasar clustering from the Two-Degree Field (2dF) and Sloan quasar surveys (see, e.g., Boyle et al. 1999; Fan et al. 1999; York et al. 2000). These measurements can test various hypotheses about the origin of quasar activity, including our primary assumption of a monotonic relation between quasar luminosity and host halo mass. If this assumption proves valid, then the first major physical result to emerge from the 2dF and Sloan measurements of high-redshift quasar clustering will be a new determination of the typical quasar lifetime.

2. METHOD

2.1. Overview

We adopt a simple model of the high-redshift quasar population that is, doubtless, idealized, but which should be reasonably accurate for our purpose of computing clustering strength as a function of quasar lifetime. We assume that all quasars reside in dark matter halos and that a given halo hosts at most one active quasar at a time. The first assumption is highly probable, since a dark matter collapse is necessary to seed the growth of a black hole, and the second should be a fair approximation at high redshift, where the masses of large halos are comparable to the masses of individual galaxy halos today.

Our strongest and most important assumption is that the luminosity of a quasar during its active phase is monotoni-

cally related to the mass of its host dark matter halo, and that all sufficiently massive halos host an active quasar at some point. More precisely, we assume that an absolute-magnitude–limited sample of quasars at redshift z samples the most massive halos present at that redshift, and that the probability that a halo above the minimum host mass $M_{\min}$ harbors an active quasar at any given time is t_Q/t_H, where t_Q is the average quasar lifetime and t_H is the halo lifetime. We can therefore compute the value of $M_{\min}$ for a quasar population with comoving space density $\Phi(z)$ from the condition

$$\Phi(z) = \int_{M_{\min}}^{\infty} dM n(M) \frac{t_Q}{t_H} . \qquad (1)$$

We compute $n(M)$ using the PS approximation, and we compute the bias of halos with $M > M_{\min}$ using the M-W approximation.

A connection between quasar luminosity and host halo mass is plausible on theoretical grounds: the cores of massive halos collapse early, giving black holes time to grow, and these halos provide larger gas supplies for fueling activity. It is also plausible on empirical grounds: local black hole masses are correlated with the host spheroid luminosity (Magorrian et al. 1998; van der Marel 1999; Salucci et al. 1999), which in turn is correlated with stellar velocity dispersion (Faber & Jackson 1976). A precisely monotonic relation is certainly an idealization, and we explore the effects of relaxing this assumption in § 4.1. The assumption of an approximately monotonic relation can be tested empirically by searching for the predicted relation between clustering strength and luminosity, as we discuss in § 4.3.

The ubiquity of black holes in luminous local spheroids supports our assumption that all sufficiently massive halos go through a quasar phase. However, once the quasar space density declines at $z < 2$, the occurrence of quasar activity must be determined by fueling rather than by the mere existence of a massive black hole, so it is not plausible that all large halos host a low-redshift quasar. We therefore apply our model only to the high-redshift quasar population, at $z \geq 2$.

We implicitly assume that a quasar turns on at a random point in the life of its host halo. In this sense, our model differs subtly from that of Haehnelt et al. (1998), who assume that a quasar turns on when the halo is formed, but this difference is unlikely to have a significant effect on the predicted clustering. Haehnelt et al. (1998) pointed out that a longer quasar lifetime would correspond to stronger quasar clustering because of the association with rarer peaks of the mass distribution, but they did not calculate this relation in detail.

Because the quasar lifetime enters our calculation only through the probability t_Q/t_H that a halo hosts an active quasar at a given time, it makes no difference whether the quasar shines continuously or turns on and off repeatedly with a short duty cycle (as argued recently by Ciotti & Ostriker 1999). For our purposes, t_Q is the total time that the quasar shines at close to its peak luminosity. We also assume that quasars radiate isotropically, with a beaming factor $f_B = 1$, but because a smaller beaming factor simply changes the conversion between observed surface density and intrinsic comoving space density, all our results can be scaled to smaller average beaming factors by replacing t_Q with $f_B t_Q$.

2.2. Notation

All our calculations assume Gaussian primordial fluctuations. We denote by $P(k)$ the power spectrum of these fluctuations as extrapolated to the present day ($z = 0$) by linear theory. The rms fluctuation of the linear density field on mass scale M is

$$\sigma(M) = \left[\frac{1}{2\pi^2} \int_0^\infty dk k^2 P(k) \tilde{W}^2(kr)\right]^{1/2}, \quad (2)$$

where

$$\tilde{W}(kr) = \frac{3(kr \sin kr - \cos kr)}{(kr)^3}, \quad r = \left(\frac{3M}{4\pi\rho_0}\right)^{1/3}, \quad (3)$$

is the Fourier transform of a spherical top-hat containing average mass M. The mean density of the universe at $z = 0$ is $\rho_0 = 2.78 \times 10^{11} \Omega_M h^2 \ M_\odot \ \text{Mpc}^{-3}$, with $h \equiv H_0/(100 \ \text{km s}^{-1} \ \text{Mpc}^{-1})$. The rms fluctuation can be considered as a function of either the mass scale M or the equivalent radius r. We define the normalization of the power spectrum by the value of $\sigma_8 \equiv \sigma(r = 8 \ h^{-1} \ \text{Mpc})$.

The rms fluctuation of the linear density field at redshift z is

$$\sigma(M, z) = \sigma(M)D(z), \quad (4)$$

where $D(z)$ is the linear growth factor $D(z)$, defined such that $D(z = 0) = 1$. The general expression for the growth factor in terms of the scaled expansion factor $y = (1 + z)^{-1}$ is

$$\delta(y) = \frac{5}{2} \Omega_M \frac{1}{y} \frac{dy}{d\tau} \int_0^y \left(\frac{dy'}{d\tau}\right)^{-3} dy', \quad (5)$$

where $D(y) = \delta(y)$ for $\Omega_M = 1$, $D(y) = \delta(y)/\delta(1)$ for $\Omega_M < 1$, and the dimensionless time variable is $\tau = H_0 t$ (Heath 1977; Carroll, Press, & Turner 1992). If the dominant energy components are pressureless matter and a cosmological constant with $\Omega_\Lambda = \Lambda/3H_0^2$, then the Friedmann equation implies

$$\left(\frac{dy}{d\tau}\right)^2 = 1 + \Omega_M(y^{-1} - 1) + \Omega_\Lambda(y^2 - 1). \quad (6)$$

For an $\Omega_M = 1$, $\Omega_\Lambda = 0$ universe, $D(z) = (1 + z)^{-1}$. Peebles (1980, eq. [11.16]) gives an exact analytic expression for $D(z)$ for the case $\Omega_M < 1$, $\Omega_\Lambda = 0$, and Carroll et al. (1992, eq. [29]) give an accurate analytic approximation for $\Omega_\Lambda \neq 0$. In our notation, $\sigma(M)$ without any explicit z always refers to the rms linear mass fluctuation on scale M at $z = 0$.

At any redshift, we can define a characteristic mass $M_*(z)$ by the condition

$$\sigma[M_*(z)] = \delta_c(z) = \frac{\delta_{c,0}}{D(z)}, \quad (7)$$

where $\delta_c(z)$ is the threshold density for collapse of a homogeneous spherical perturbation at redshift z. Because we implicitly define the density field as "existing" at $z = 0$, the collapse threshold $\delta_c(z)$ increases with increasing redshift. For an $\Omega_M = 1$, $\Omega_\Lambda = 0$ universe, $\delta_{c,0} = 0.15(12\pi)^{2/3} \approx 1.69$ (see, e.g., Peebles 1980, § 19). For other models, we incorporate the dependence of $\delta_{c,0}$ on Ω_M in Appendix A of Navarro, Frenk, & White (1997), but because Ω_M approaches one at high redshift in all models, this correction to δ_c is less than 2% in all the cases that we consider.

2.3. From the Quasar Lifetime to the Minimum Halo Mass

For a specified quasar lifetime, we compute the minimum halo mass by matching the comoving number density, $\Phi(z)$, of observed quasars, accounting for the fact that only a fraction t_Q/t_H of host halos will have an active quasar at the time of observation. The matching condition is equation (1), or, putting in explicit mass and redshift dependences,

$$\Phi(z) = \int_{M_{\min}}^\infty dM \frac{t_Q}{t_H(M, z)} n(M, z). \quad (8)$$

If $t_Q > t_H(M, z)$, we set the factor t_Q/t_H to unity. For the halo number density, we use the P-S approximation,

$$n(M, z)dM = -\sqrt{\frac{2}{\pi}} \frac{\rho_0}{M} \frac{\delta_c(z)}{\sigma^2(M)} \frac{d\sigma(M)}{dM}$$

$$\times \exp\left[-\frac{\delta_c^2(z)}{2\sigma^2(M)}\right]dM, \quad (9)$$

where ρ_0 is the mean density of the universe at $z = 0$, $\sigma(M)$ is the rms fluctuation given by equation (2), and $\delta_c(z)$ is the critical density for collapse by redshift z.

In a gravitational clustering model of structure formation, halos are constantly growing by accretion and mergers, so the definition of a "halo lifetime" is somewhat ambiguous. For $\Omega_M(z) \approx 1$, a typical halo survives for roughly a Hubble time before being incorporated into a substantially larger halo, since the age of the universe at redshift z is also the characteristic dynamical time of objects forming at that redshift. Thus, to a first approximation, one could simply substitute $t_H(M, z) = t_U(z)$ in equation (9). We can do somewhat better by using the extended Press-Schechter formalism (e.g., Bond et al. 1991; Lacey & Cole 1993) to calculate the average halo lifetime, thereby accounting for the dependence of t_H on the power spectrum shape and the halo mass. Structure grows more rapidly in a cosmology with a redder power spectrum, and more massive halos accrete mass more rapidly.

Equation (2.22) of Lacey & Cole (1993) gives the probability that a halo of mass M_1 existing at time t_1 will have been incorporated into a new halo of mass greater than M_2 by time t_2:

$$P(S < S_2, \omega_2 | S_1, \omega_1) = \frac{1}{2} \frac{(\omega_1 - 2\omega_2)}{\omega_1} \exp\left[\frac{2\omega_2(\omega_1 - \omega_2)}{S_1}\right]$$

$$\times \left\{1 - \text{erf}\left[\frac{S_2(\omega_1 - 2\omega_2) + S_1 \omega_2}{\sqrt{2S_1 S_2(S_1 - S_2)}}\right]\right\}$$

$$+ \frac{1}{2}\left\{1 - \text{erf}\left[\frac{S_1 \omega_2 - S_2 \omega_1}{\sqrt{2S_1 S_2(S_1 - S_2)}}\right]\right\}, \quad (10)$$

where $S_1 = \sigma^2(M_1)$, $S_2 = \sigma^2(M_2)$, $\omega_1 = \delta_c(t_1)$, and $\omega_2 = \delta_c(t_2)$. In this equation, ω plays the role of the "time" variable, with $\omega_2 < \omega_1$ corresponding to $t_2 > t_1$, and S plays the role of the "mass" variable, with $S_2 < S_1$ corresponding to $M_2 > M_1$. For a halo of mass M existing at time $t_U(z)$, we define the halo lifetime to be the median interval before such a halo is incorporated into a halo of mass $2M$. Thus, $t_H(M, z) = \hat{t}_S - t_U(z)$, where $\hat{t}_S$ is the time at which the probability in equation (10) equals 0.5, for $S_1 = \sigma^2(M)$ and $S_2 = \sigma^2(2M)$. Clearly, other plausible definitions of $t_H(M, z)$ are possible, and they would give answers different by factors of order unity. With our definition, a black hole that lights up

repeatedly is considered the "same" quasar as long as the mass of its host halo remains the same within a factor of 2. If the host merges into a much larger halo and the black hole lights up again, it is considered a "new" quasar. We show the halo lifetimes for different masses and power spectra when we discuss specific models below.

For comoving space densities $\Phi(z)$, we adopt values based on the work of Boyle et al. (1990), Hewett, Foltz, & Chaffee (1993), and Warren, Hewett, & Osmer (1994). Since observations constrain the number of objects per unit redshift per unit solid angle, the conversion to comoving space density depends on the values of the cosmological parameters. We provide the formulae for these conversions in the Appendix, and in Table A1 we list our adopted values of $\Phi(z)$ and the surface densities of objects to which these space densities correspond. In general, $\Phi(z)$ represents the space density of quasars above some absolute magnitude, corresponding to a surface density above some apparent magnitude. In § 4.3 we discuss how to scale our results to predict the clustering of samples with different measured surface densities.

2.4. From Minimum Halo Mass to Clustering Length

Halos with $M > M_*$ are clustered more strongly than the underlying distribution of mass. MW give an approximate formula,

$$b(M, z) = 1 + \frac{1}{\delta_{c,0}}\left[\frac{\delta_c^2(z)}{\sigma^2(M)} - 1\right], \qquad (11)$$

for the bias factor of halos of mass M at redshift z. On large scales, the ratio of rms fluctuations in halo number density to rms fluctuations in mass should be $b(M, z)$. This formula is derived from an extended Press-Schechter analysis, and it agrees fairly well with the results of N-body simulations on scales where the rms mass fluctuations are less than unity. The MW formula becomes less accurate for halos with $M < M_*$, i.e., $\sigma(M) < \delta_c(z)$, and Jing (1998) provides an empirical correction that fits the N-body results,

$$b(M, z) = \left\{1 + \frac{1}{\delta_{c,0}}\left[\frac{\delta_c^2(z)}{\sigma^2(M)} - 1\right]\right\} \\ \times \left[\frac{\sigma^4(M)}{2\delta_c^4(z)} + 1\right]^{(0.06 - 0.02n_{\mathrm{eff}})}, \qquad (12)$$

where $n_{\mathrm{eff}} = 3 - 6(d \ln \sigma / d \ln M)$ is the effective index of the power spectrum on a mass scale M.

According to our model, the quasars at redshift z only reside in halos of mass $M > M_{\min}$. The effective bias of these host halos is the bias factor (eq. [12]) weighted by the number density and lifetime of the corresponding halos:

$$b_{\mathrm{eff}}(M_{\min}, z) = \int_{M_{\min}}^{\infty} dM\, \frac{b(M, z)n(M, z)}{t_H(M, z)} \\ \times \left[\int_{M_{\min}}^{\infty} dM\, \frac{n(M, z)}{t_H(M, z)}\right]^{-1}. \qquad (13)$$

Because the halo number density drops steeply with increasing mass, the effective bias is usually only slightly larger than the bias factor at the minimum halo mass, $b(M_{\min}, z)$.

As our measure of clustering amplitude, we use the radius r_1 of a top-hat sphere in which the rms fluctuation σ_Q of quasar number counts (in excess of Poisson fluctuations) is

unity. This quantity is similar to the correlation length r_0 at which the quasar correlation function $\xi(r)$ is unity, but it can be more robustly constrained observationally because it does not require fitting the scale-dependence of $\xi(r)$. For a power-law correlation function $\xi(r) = (r/r_0)^{-1.8}$, $r_1 \approx 1.4r_0$. With our adopted approximation for the bias, r_1 is determined implicitly by the condition

$$\sigma_Q(r_1, z) = b_{\mathrm{eff}}(M_{\min}, z)\sigma(r_1)D(z) = 1, \qquad (14)$$

where $\sigma(r_1)$ is the rms linear mass fluctuation at $z = 0$ in spheres of radius r_1. For a specified cosmology, mass power spectrum $P(k)$, quasar lifetime t_Q, and comoving space density $\Phi(z)$, we determine r_1 from equation (14), computing $\sigma(r)$ from equation (2), $D(z)$ from equation (5), $M_{\min}$ from equation (8), and $b_{\mathrm{eff}}(M_{\min})$ from equations (12) and (13).

2.5. Results for Power-Law Power Spectra

Models with a power-law power spectrum, $P(k) \propto k^n$, provide a useful illustration of our methods, since many steps of the calculation can be done analytically. For such models, the dependence of rms fluctuation on mass is also a power law,

$$\sigma(M) = \frac{\sigma(M, z)}{D(z)} = \frac{\delta_{c,0}}{D(z)}\left[\frac{M}{M_{*(z)}}\right]^{-(3+n)/6}$$

$$= \delta_c(z)\left[\frac{M}{M_{*(z)}}\right]^{-(3+n)/6}, \qquad (15)$$

where $M_*(z)$ is the characteristic nonlinear mass defined by equation (7). With this substitution, the PS mass function can be expressed as a function of $M_*(z)$ and the dimensionless mass variable $x = M/M_*(z)$. Integrating to obtain the comoving number density of objects with mass $M > M_{\min}$ yields

$$N(M > M_{\min}) = \sqrt{\frac{2}{\pi}}\left(\frac{n+3}{6}\right)\left(\frac{\rho_0}{M_*}\right)$$

$$\times \int_{x_{\min}}^{\infty} dx\, x^{(n-9)/6} \exp\left[-\frac{1}{2}x^{(n+3)/3}\right]. \qquad (16)$$

For power-law models with $\Omega_M = 1$, the ratio of the halo lifetime $t_H(M, z)$ to the age of the universe at redshift z depends only on $M/M_*(z)$ and has no separate dependence on redshift. Figure 1 shows t_H and the ratio t_H/t_U as a function of M/M_* for power-law models with $n = 0, -1$, and -2. More massive halos tend to accrete mass more quickly and therefore have shorter median lifetimes. At a given value of M/M_*, the halo lifetime is shorter for lower n because a greater amount of large-scale power causes the typical mass scale of nonlinear structure to grow more rapidly. Although the calculation of the median halo lifetime via equation (10) is moderately complicated, the median lifetime for large masses asymptotically approaches a constant value

$$t_H(M, z) = [2^{(3+n)/2} - 1]t_U(z), \quad M \gg M_*(z), \Omega_M = 1, \qquad (17)$$

(Lacey & Cole 1993). We find below that the predicted masses of quasar host halos are indeed in this asymptotic regime for the most plausible parameter choices. The halo lifetime is longer for $\Omega_M < 1$ than for $\Omega_M = 1$ because fluctuations grow more slowly in a low-density universe, but t_H still asymptotically approaches a constant value. The dotted curves in Figure 1 illustrate the case of $n = -1$,
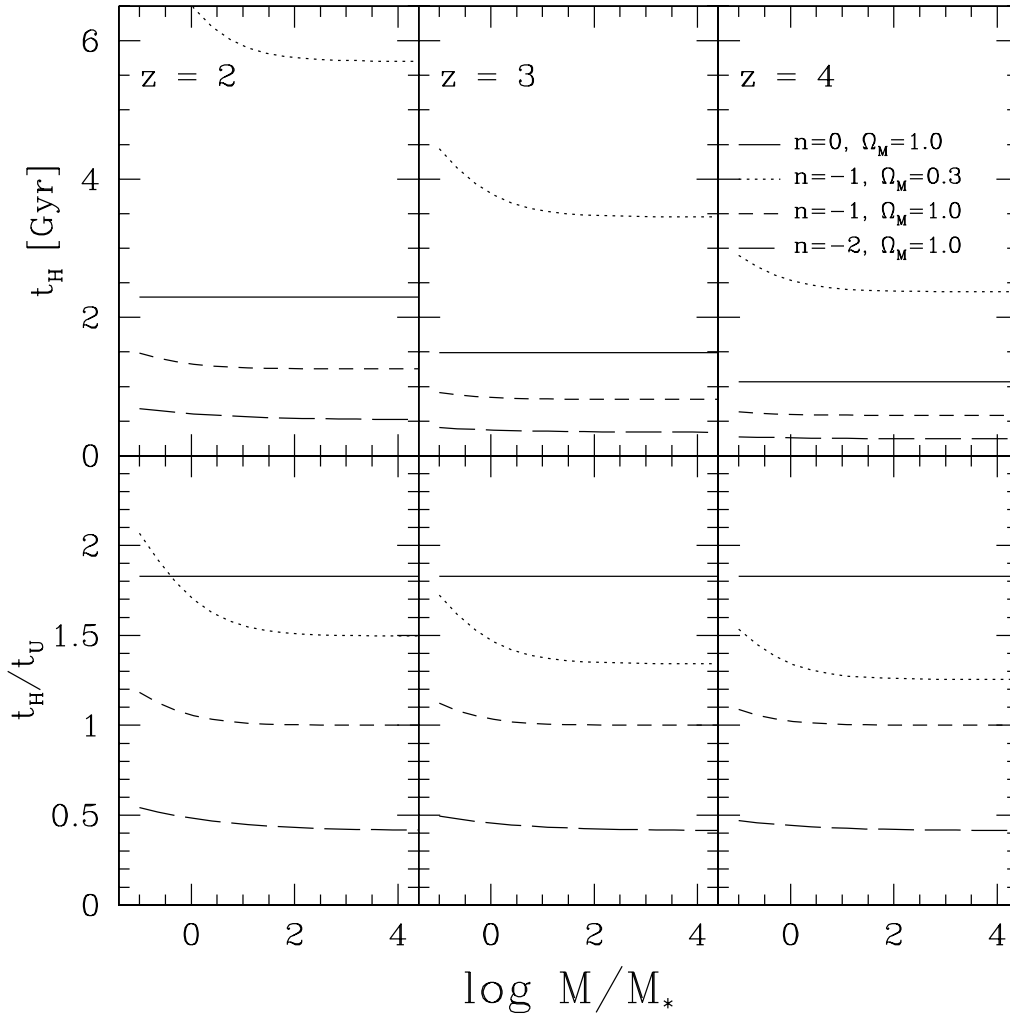
FIG. 1.—Halo lifetimes vs. M/M_* for the power-law cosmologies listed in Table 1. Top panels show the halo lifetimes in Gyr for each model at $z = 2$, 3, and 4. Bottom panels show the ratio of the halo lifetime to the age of the universe. This ratio is independent of redshift for the $\Omega_M = 1$ models, but not for the $n = -1, \Omega_M = 0.3$ model.

$\Omega_M = 0.3$, $\Omega_\Lambda = 0$. The cosmological parameters for all our models with power-law power spectra are summarized in Table 1.

The power-law scaling of the rms fluctuation amplitude, equation (15), allows the bias equation (12) to be written

$$b(M, z) = \left(1 + \frac{1}{\delta_{c,0}} \left\{\left[\frac{M}{M_*(z)}\right]^{(3+n)/3} - 1\right\}\right)$$
$$\times \left\{\frac{1}{2}\left[\frac{M}{M_*(z)}\right]^{-2(3+n)/3} + 1\right\}^{(0.06-0.02n)}. \quad (18)$$

Note that the second factor is very close to 1 for $M \geq M_*$. Figure 2 shows the bias and the corresponding number-weighted effective bias (eq. [13]) as a function of $M_{\min}/M_*$. For $M_{\min} > M_*$, the effective bias is only slightly larger than $b(M_{\min})$, since the number density of halos declines rapidly with increasing M. As equation (18) shows, the bias depends more strongly on M for larger n. However, the exponentially falling tail of the mass function at high M/M_* is much steeper for higher n, as one can see from equation (16). As a result, the bias at fixed comoving number density is higher for *smaller* n in the high M/M_* regime (see Fig. 3 below).

Under the (good) approximation that the halo lifetime is given by the asymptotic formula (eq. [17]) in the mass range of interest, the halo lifetime can be moved outside the integral of equation (8) for the number density of active quasars. The implied quasar lifetime as a function of minimum halo mass is then

$$t_Q(M_{\min}) = \frac{t_H \Phi(z)}{N(M > M_{\min})}, \quad (19)$$

where $N(M > M_{\min})$ is given by equation (16). For the $\Omega_M = 0.3$, $n = -1$ model, we also use the asymptotic value of t_H, although this is no longer given by equation (17). We use a $P(k)$ normalization $\sigma_8 = 0.5$ for the three $\Omega_M = 1$ models and $\sigma_8 = 1.0$ for the $\Omega_M = 0.3$ model, in approximate agreement with the constraint on σ_8 and Ω_M implied by the observed mass function of rich galaxy clusters (White, Efstathiou, & Frenk 1993; Eke, Cole, & Frenk 1996).

Equation (19) implicitly determines $M_{\min}/M_*(z)$ given t_Q. The top panels of Figure 3 show $M_{\min}/M_*(z)$ as a function of t_Q/t_U for $z = 2$, 3 and 4 and a constant comoving space density $\Phi(z) = 10^{-6}$ h^3 Mpc^{-3}. For the $\Omega_M = 1$ cases, where t_Q/t_U depends only on n and $M_{\min}/M_*$, the redshift dependence of $M_{\min}/M_*$ arises solely from the presence of

TABLE 1

POWER-LAW MODEL PARAMETERS

n (1)	σ_8 (2)	h (3)	Ω_M (4)	$M_*(z=2)$ (5)	$M_*(z=3)$ (6)	$M_*(z=4)$ (7)
0......	0.5	1.0	1.0	5.87×10^{12}	3.30×10^{12}	2.11×10^{12}
-1......	0.5	1.0	1.0	5.80×10^{12}	2.45×10^{12}	1.25×10^{12}
-1......	1.0	1.0	0.3	5.44×10^{12}	2.87×10^{12}	1.70×10^{12}
-2......	0.5	1.0	1.0	5.59×10^{8}	9.96×10^{7}	2.61×10^{7}

NOTE.—Parameters of the four power-law cosmological models discussed in § 2.5. Cols. (1)–(4) list the power spectrum index n, normalization σ_8, scaled Hubble constant h, and mass density parameters Ω_M. Cols. (5)–(7) list the values of M_* (eq. [7]) at $z = 2, 3,$ and 4, in units of $h^{-1}\ M_\odot$.

M_* in the number density equation (16). As M_* increases with decreasing redshift, the value of $x_{\min} = M_{\min}/M_*$ must decrease to keep $\Phi(z)$ constant. Smaller values of n lead to higher values of $M_{\min}/M_*$ because of the gentler falloff of the mass function at large M/M_*. The $n = -2$ curves become flat at the largest values of t_Q/t_U because t_Q begins to exceed the halo lifetime t_H, implying that all halos above $M_{\min}$ are occupied by quasars. The difference between the open and $\Omega_M = 1$ curves for $n = -1$ reflects mainly the larger values of σ_8 and $D(z)$ in the open model, which lead to a lower value of ρ_0/M_* in the mass function (eq. [16]) and therefore require a lower value of $M_{\min}/M_*$ to compensate.

The middle panels of Figure 3 show the effective bias, $b_{\mathrm{eff}}(M_{\min}, z)$, for the power-law models. As already remarked, the bias at fixed space density and t_Q/t_U is higher for redder power spectra (smaller n) because of the much higher values of $M_{\min}/M_*$, despite the partially counterbalancing effect of the stronger dependence of bias on mass at larger n. Physically, the higher bias for redder spectra reflects the greater influence of the large-scale environment on the amplitude of small-scale fluctuations. For a given model, the bias increases with increasing redshift, reflecting the increase in $M_{\min}/M_*$; the change, however, is quite modest.

The rms number count fluctuation on comoving scale r is

$$\sigma_Q(r, z) = b_{\mathrm{eff}}(M_{\min}, z)\sigma(r, z)$$
$$= b_{\mathrm{eff}}(M_{\min}, z)\sigma_8 D(z)\left(\frac{r}{8\ h^{-1}\ \mathrm{Mpc}}\right)^{-(3+n)/2}. \quad (20)$$

The quasar clustering length r_1 is the scale on which this rms fluctuation amplitude is unity,

$$r_1 = (8\ h^{-1}\ \mathrm{Mpc})[b_{\mathrm{eff}}(M_{\min}, z)\sigma_8 D(z)]^{2/(3+n)}. \quad (21)$$

The bottom panels of Figure 3 present the main result of this section, the dependence of r_1 on quasar lifetime for our four power-law models at $z = 2, 3,$ and 4. As anticipated, the quasar clustering length shows a strong dependence on quasar lifetime. The relation between r_1 and t_Q depends on the power spectrum index n, so the shape of the power spectrum must be known fairly well to determine t_Q from measurements of r_1. The clustering at fixed t_Q/t_U is substantially stronger in the open $n = -1$ model than in the $\Omega_M = 1$ model, because the underlying mass distribution is more strongly clustered [larger σ_8 and $D(z)$].

For a specified value of Ω_M, the cluster mass function imposes a reasonably tight constraint on the normalization, σ_8. It is nonetheless interesting to explore the sensitivity of the predicted quasar clustering to this normalization. More intuitive than the σ_8-dependence is the equivalent relation between the quasar clustering length and the corresponding clustering length of the underlying mass distribution at the same redshift,

$$r_{1m} = (8\ h^{-1}\ \mathrm{Mpc})[\sigma_8 D(z)]^{2/(3+n)}. \quad (22)$$

Figure 4 plots this relation at $z = 3$ for the four power-law cosmologies and $t_Q = t_U$ (top curve), $0.1t_U$, $0.01t_U$, and $0.001t_U$ (bottom curve), for values of σ_8 ranging from 0.2 to 2.0. The values of r_{1m} that correspond to the σ_8 values in Table 1 are marked with open circles. If the bias did not change with r_{1m}, then the quasar clustering length r_1 would grow in proportion to r_{1m}, and the curves in Figure 4 would parallel the diagonal of the box, which has a slope of 1.0. However, increasing r_{1m} increases M_*, and therefore requires a lower value of $M_{\min}/M_*$ to match the quasar space density. The correspondingly lower bias partially compensates for the larger r_{1m}, making the curves in Figure 4 shallower than the box diagonal. For $n = 0$ and large t_Q/t_U (the highest solid curve), the minimum mass $M_{\min}$ lies far out on the tail of a steeply falling mass function. In this regime, a change in M_* requires only a small change in
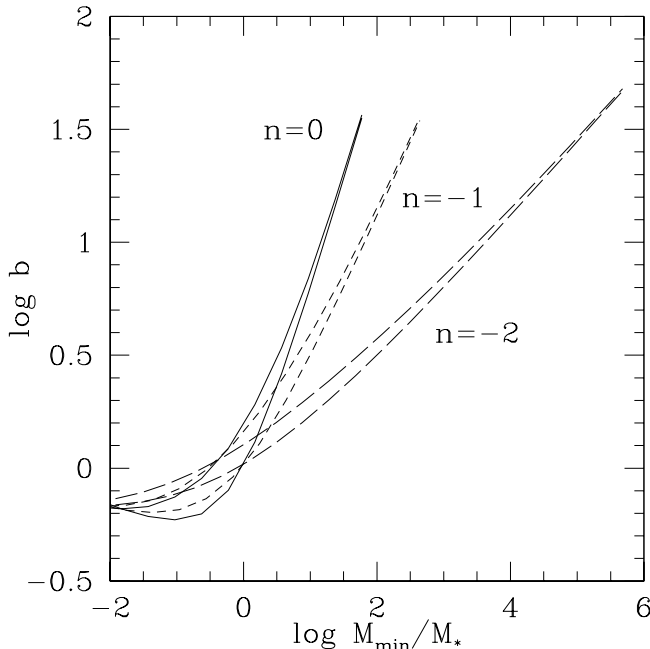

FIG. 2.—Bias vs. M/M_* for the power-law models with $n = 0$ (*solid line*), -1 (*short-dashed line*), and -2 (*long-dashed line*). Lower curves show $b(M_{\min})$ computed from eq. (12), and upper curves show the number-weighted effective bias (eq. [13]).
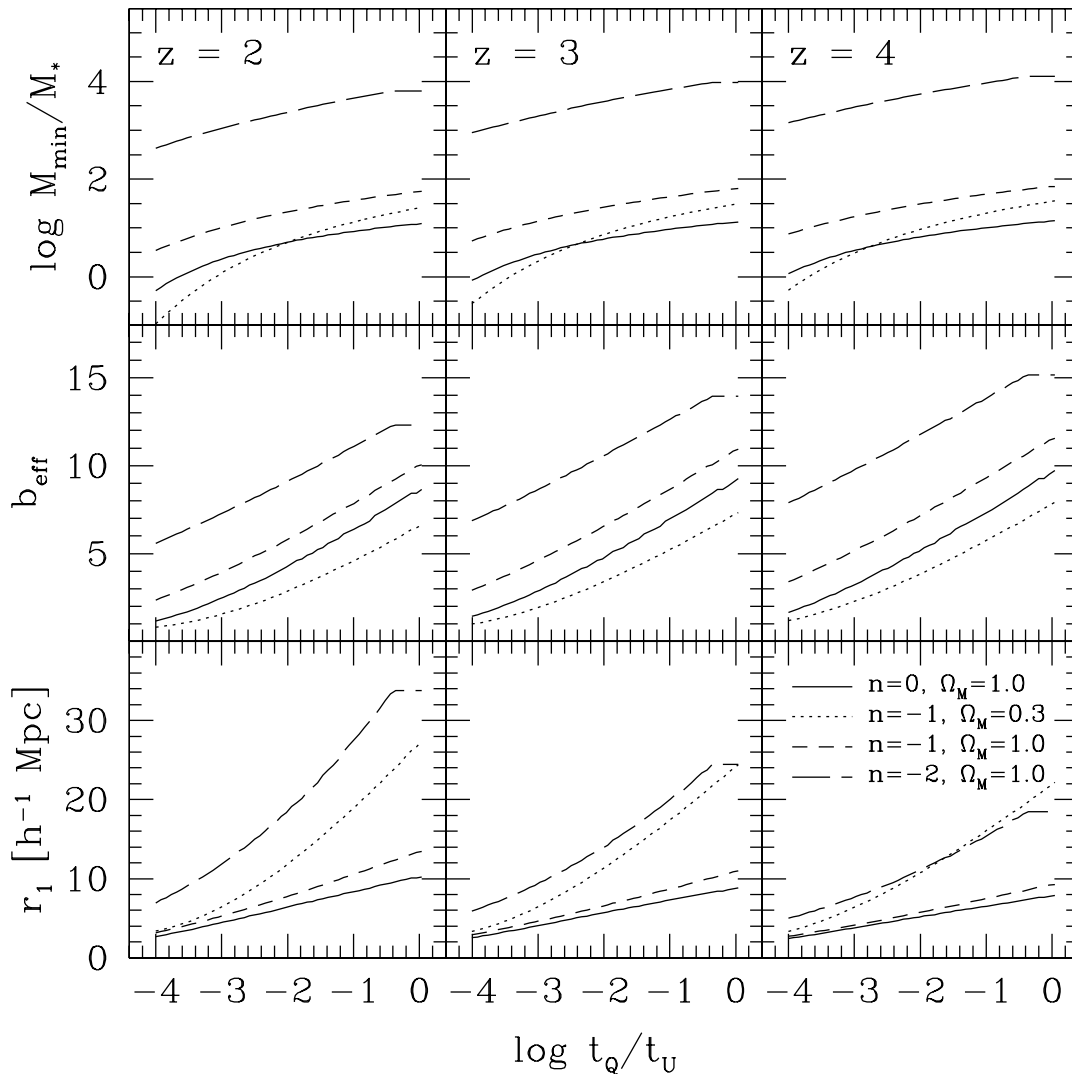
Fig. 3.—Halo mass, bias, and clustering length for the power-law models, as a function of quasar lifetime. Top panels show the minimum mass, M_{min}, required to obtain a space density of $\Phi(z) = 10^{-6} \, h^3 \, \mathrm{Mpc}^{-3}$ for a given value of t_Q/t_U at $z = 2, 3$, and 4. Middle panels show the corresponding effective bias, b_{eff}, for halos M_{min}. Bottom panels show the clustering lengths, r_1, as a function of t_Q/t_U. The clustering length is the radius of a top-hat sphere in which rms number count fluctuations (in excess of Poisson) are unity. For a power-law correlation function $\xi(r) = (r/r_0)^{-1.8}$, $r_1 \approx 1.4 r_0$.

M_{min}/M_* to compensate, so there is little change in b_{eff} with r_{1m}, and the curves approach the $r_1 \propto r_{1m}$ lines that would apply for constant bias. A similar argument explains the steepening of all curves toward low r_{1m}, where the small values of M_* put the value of M_{min} farther out on the tail of the mass function.

3. RESULTS FOR COLD DARK MATTER COSMOLOGIES

The results of § 2.5 confirm our initial contention that quasar clustering can provide a good diagnostic of the typical quasar lifetime. However, they show that the predicted clustering length also depends on the shape of the mass power spectrum and on the value of Ω_M, which influences the cluster normalization of σ_8 at $z = 0$ and (together with Ω_Λ) determines the growth factor $D(z)$. Accurate determination of t_Q from measurements of quasar clustering therefore requires reasonably good knowledge of the underlying cosmology. Fortunately, many lines of evidence now point toward a flat, low-density model based on inflation and cold dark matter (see, e.g., the review by Bahcall et al. 1999). In particular, recent studies of the power spectrum of

the Lyα forest imply that the matter power spectrum has the shape and amplitude predicted by COBE- and cluster-normalized cold dark matter (CDM) models with $\Omega_M \sim 0.4$ at the redshifts and length scales relevant to the prediction of quasar clustering (Croft et al. 1999; Weinberg et al. 1999; McDonald et al. 2000; Phillips et al. 2000).

For the power spectrum of our CDM models, we adopt $P(k) \propto k^{n_p} T^2(k)$ with scale-invariant ($n_p = 1$) primeval inflationary fluctuations and the transfer function parameterization of Bardeen et al. (1986),

$$T(k) = \frac{\ln(1 + 2.34q)}{2.34q} \left[1 + 3.89q + (16.1q)^2 \right.$$
$$\left. + (5.46q)^3 + (6.71q)^4\right]^{-1/4} . \qquad (23)$$

Here $q = k/\Gamma$ and Γ, with units of $(h^{-1} \, \mathrm{Mpc})^{-1}$, is the CDM shape parameter, given approximately by $\Gamma = \Omega_M h \exp\left[-\Omega_b - (2h)^{1/2}\Omega_b/\Omega_M\right]$ (Sugiyama 1995). We calculate $\sigma(M)$ and $(d\sigma/dM)$ by numerical integration of this power spectrum.

We consider five different CDM models with the parameters listed in Table 2. These models are chosen to illustrate a
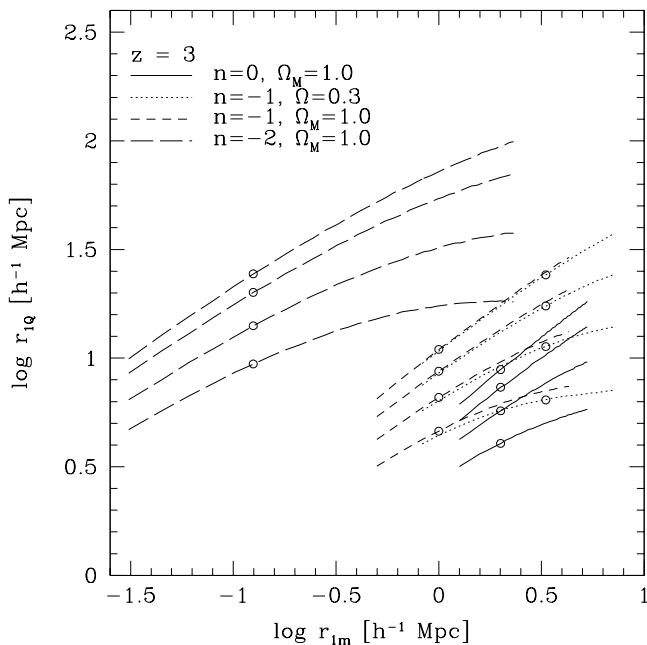
FIG. 4.—Dependence of the quasar clustering length, r_{1Q}, on the mass clustering length, r_{1m}, for the four power-law models at $z = 3$. In each case, the four lines show, from top to bottom, the lifetimes $t_Q = t_U$, $0.1t_U$, $0.01t_U$, and $0.001t_U$. Results are computed for normalizations running from $\sigma_8 = 0.2$ to $\sigma_8 = 2.0$. Open circles show r_{1m} and r_{1Q} for our standard choices of σ_8, listed in Table 1. If bias were independent of r_{1m}, the lines would parallel the diagonal of the box, which has a slope of 1.0.

range of cosmological inputs and also to isolate the effects of different parameters on quasar clustering predictions. The τCDM, OCDM, and ΛCDM models have $\Gamma = 0.2$, in approximate agreement with the shape parameter estimated from galaxy surveys (e.g., Baugh & Efstathiou 1993; Peacock & Dodds 1994), and they have σ_8 values consistent with the cluster mass function constraints of Eke et al. (1996). The τCDM and ΛCDM models are approximately COBE-normalized. COBE normalization would imply a lower σ_8 for OCDM, but a slight increase in n_p could raise σ_8 without having a large impact on the shape of $P(k)$ at the relevant scales. The OCDM and ΛCDM models are consistent with the Lyα forest power spectrum measurements of Croft et al. (1999), but the τCDM model is not. OCDM is inconsistent with the observed location of the first acoustic peak in the cosmic microwave background anisotropy spectrum (e.g., Miller et al. 1999; Melchiorri et al. 2000; Tegmark & Zaldarriaga 2000), and of the three models, only ΛCDM is consistent with the Hubble diagram of Type Ia supernovae (Riess et al. 1998; Perlmutter et al. 1999).

We use the comparison between the τCDM and SCDM models, with $\Gamma = 0.2$ and $\Gamma = 0.5$, respectively, to illustrate the impact of the power-spectrum shape for fixed Ω_M and σ_8. The SCDM model is cluster-normalized, but its $\sigma_8 = 0.5$ is well below the value of $\sigma_8 \approx 1.2$ implied by COBE for $n_p = 1$, $\Gamma = 0.5$ (e.g., Bunn & White 1997). The OCDM and ΛCDM models have the same $P(k)$ shape and the same $P(k)$ amplitude at $z = 0$, but at high redshift the OCDM model has stronger fluctuations as a result of a larger $D(z)$. We therefore include the model ΛCDM2, which has σ_8 chosen to yield the same power spectrum amplitude as OCDM at $z = 3$. Differences between OCDM and ΛCDM2 isolate the impact of a cosmological constant for fixed high-redshift mass clustering.

Figure 5 shows t_H in Gyr (top panels) and t_H/t_U (bottom panels) as a function of M/M_* for the CDM models at $z = 2$, 3, and 4. In contrast to the power-law models shown in Figure 1, the ratio t_H/t_U does not approach a constant value but instead increases at very large M/M_*. This increase can be understood with reference to the power-law case: the effective power-law index, $n_{eff} = 3 - 6(d\ln\sigma/d\ln M)$, increases with increasing mass in a CDM spectrum, and larger values of n_{eff} correspond to slower growth of mass scales (and larger t_H/t_U), as shown in Figure 1. The difference between the SCDM and τCDM curves in Figure 5 reflects the higher values of n_{eff} for the $\Gamma = 0.5$ power spectrum. The differences between the various $\Gamma = 0.2$ models largely reflect the differences in M_*, and hence differences in n_{eff} at fixed M/M_*, and they also reflect the differences in fluctuation growth rates.

Figure 6 plots the effective bias against M_{min}/M_* for the five CDM models at $z = 3$. Figure 2 showed that the value of b_{eff} at fixed M_{min}/M_* is higher for larger n. The lines in Figure 6 curve upward because n_{eff} increases with mass scale, and to a good approximation the value of b_{eff} in the CDM models equals the value of b_{eff} at the same M_{min}/M_* in a power-law model of index $n_{eff}(M_{min})$. The difference between SCDM and τCDM in Figure 6 therefore reflects the higher n_{eff} values in SCDM, and the differences among the other models reflect the different values of M_*, and hence the different values of n_{eff} at fixed M_{min}/M_*.

The top three panels of Figure 7 show the dependence of M_{min}/M_* on t_Q at $z = 2$, 3, and 4; the values of M_* are listed in Table 2. The calculation of M_{min} via equation (8) incorporates both the dependence of halo lifetime on mass and the influence of Ω_M and Ω_Λ on the value of $\Phi(z)$ inferred from the quasar surface density (as discussed in the Appendix). The two $\Omega_M = 1$ models have the lowest values of M_* because of their lower σ_8 and $D(z)$, so they require the largest M_{min}/M_* to match the observed $\Phi(z)$. The value of

TABLE 2

CDM MODEL PARAMETERS

Model (1)	σ_8 (2)	h (3)	Ω_M (4)	Ω_Λ (5)	Γ (6)	$M_*(z=2)$ (7)	$M_*(z=3)$ (8)	$M_*(z=4)$ (9)
SCDM	0.5	0.5	1.0	0.0	0.5	3.58×10^9	2.26×10^8	1.86×10^7
τCDM	0.5	0.5	1.0	0.0	0.2	9.44×10^6	1.09×10^5	1.87×10^3
OCDM.......	0.9	0.65	0.3	0.0	0.2	1.50×10^{11}	2.70×10^{10}	5.60×10^9
ΛCDM	0.9	0.65	0.3	0.7	0.2	2.70×10^{10}	2.03×10^9	1.91×10^8
ΛCDM2......	1.17	0.65	0.3	0.7	0.2	2.16×10^{11}	2.34×10^{10}	3.10×10^9

NOTE.—Parameters of the five CDM models discussed in § 3. Col. (1) lists the model name, cols. (2)–(5) the power spectrum normalization and cosmological parameters, and col. (6) the power spectrum shape parameter (see eq. [23]). Cols. (7)-(9) list the values of M_* (eq. [7]) at $z = 2$, 3, and 4, in units of $h^{-1} M_\odot$.
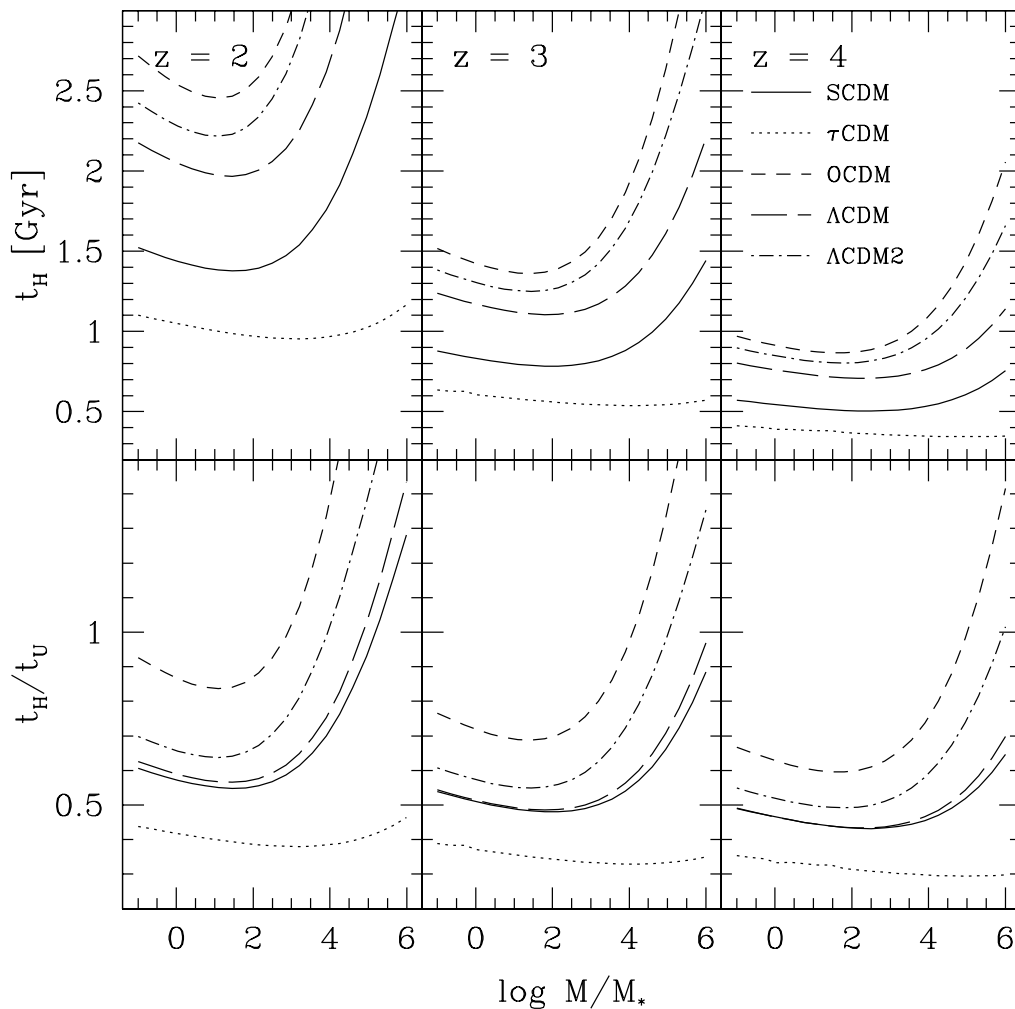
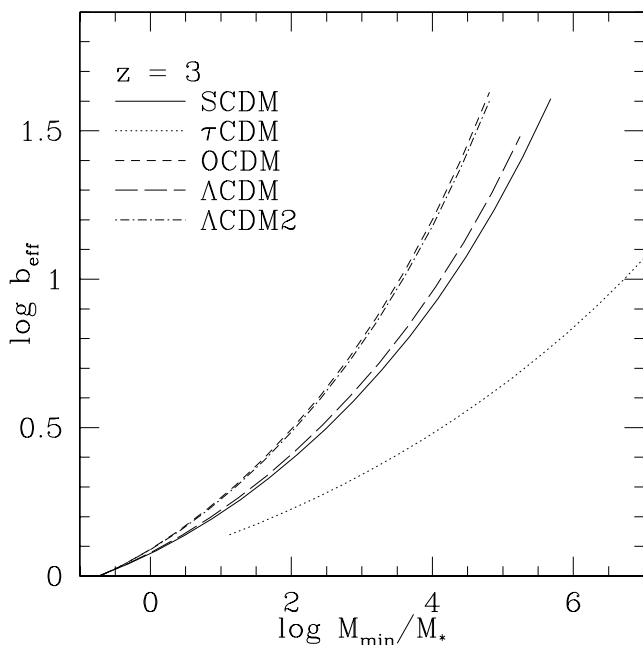FIG. 5.—Halo lifetime as a function of M/M_*, as in Fig. 1, for the CDM models with parameters listed in Table 2

FIG. 6.—Effective bias as a function of minimum halo mass, as in Fig. 2, for the CDM models at $z = 3$.

M_* is smaller for ΛCDM than for OCDM because $D(z)$ is smaller for the flat model, so ΛCDM requires larger M_{min}/M_*. The higher normalization of the ΛCDM2 model largely removes this difference, since $\sigma_8 D(z = 3)$ is matched to that of the OCDM model, but ΛCDM2 still has a slightly lower M_* because of the influence of Ω_Λ on $\delta_c(z)$. As a result, the M_{min}/M_* curve for ΛCDM2 lies just above that of OCDM at $z = 3$.

The middle panels of Figure 7 show the effective bias values, which display the same relative dependence on t_Q and cosmology as the M_{min}/M_* values. Because $M_{min}/M_* > 1$ in all the CDM models, even for t_Q as low as 10^5 yr, the MW bias formula (eq. [11]) yields nearly identical results to Jing's (1998) corrected formula (eq. [12]).

The bottom panels of Figure 7 present the main results of this paper: the relation between the clustering length r_1 and the quasar lifetime t_Q for CDM models at $z = 2$, 3, and 4. The clustering length is an increasing function of quasar lifetime for the reasons outlined in § 1 and detailed in § 2. A longer t_Q implies that quasar host halos are rarer, more highly biased objects. The change in the r_1 versus t_Q relation with redshift reflects the evolution of the quasar space density and of the underlying mass fluctuations. For a given model and t_Q, the predicted quasar clustering is weakest at
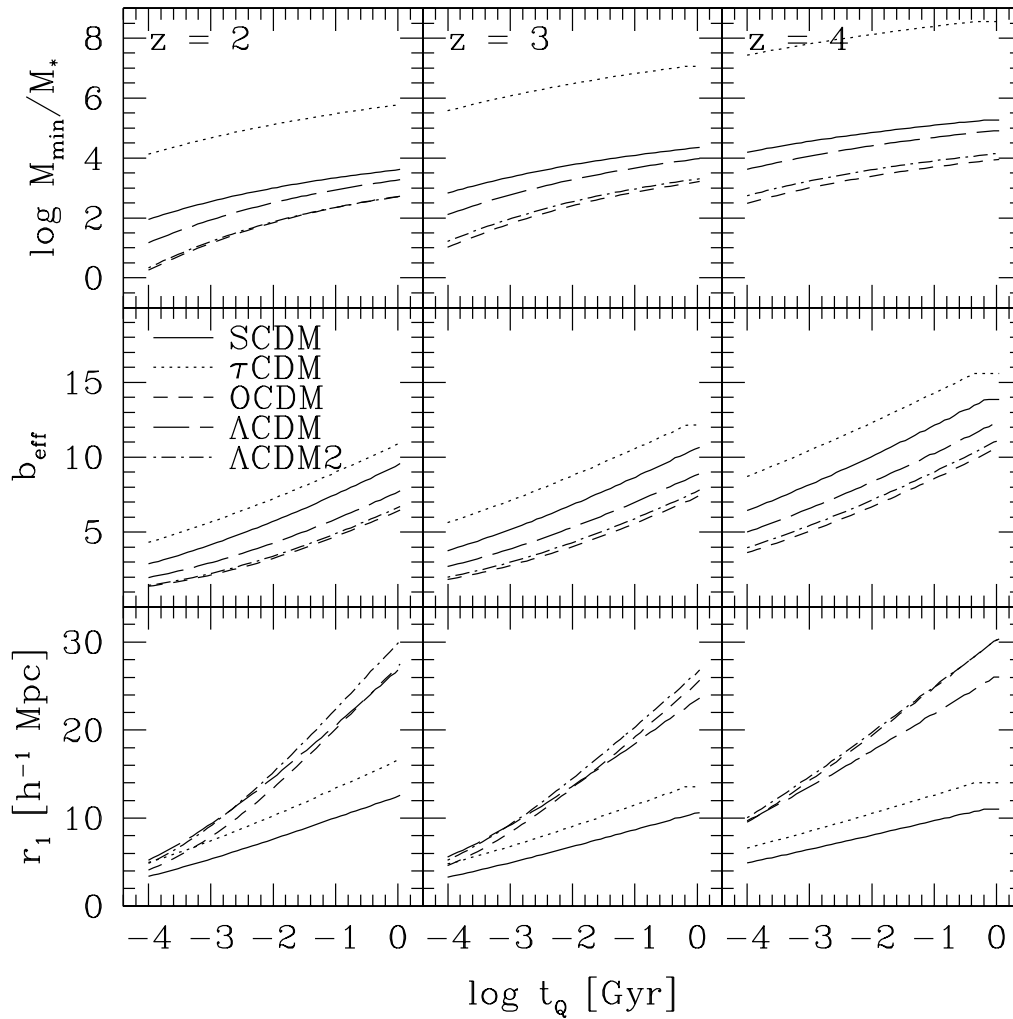
Fig. 7.—Minimum halo mass, effective bias, and clustering length as a function of t_Q, for the CDM models at $z = 2$, 3, and 4. Format is the same as in Fig. 3.

$z = 3$, the peak of the quasar space density. The smaller quasar abundance at $z = 4$ implies a higher bias of the host halo population, which more than compensates for the slightly weaker mass clustering. The clustering length grows between $z = 3$ and $z = 2$ because of both the drop in quasar space density and the growth of mass clustering. The r_1 versus t_Q relation becomes flat at the largest t_Q for the SCDM model at $z = 4$ and for the τCDM model at $z = 3$ and 4, where t_Q exceeds the halo lifetime $t_H(M_{min})$, and the

value of M_{min} required to match $\Phi(z)$ therefore becomes independent of t_Q.

The differences between models reflect the differences in bias factors discussed above and the differences in the mass clustering. There are also differences in the values of $\Phi(z)$ inferred from the observed quasar surface density (see Appendix), but these have relatively little effect. The main separation in Figure 7 is between the low-density models and the $\Omega_M = 1$ models, which have weaker mass clustering

TABLE 3

r_1 vs. t_Q Fitting Coefficients

Model	$z = 2$			$z = 3$			$z = 4$		
	a_0	a_1	a_2	a_0	a_1	a_2	a_0	a_1	a_2
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
SCDM	12.47	2.437	0.1001	10.62	1.932	0.0601	11.26	1.590	-0.0042
τCDM	16.54	3.163	0.1662	13.89	2.422	0.0882	14.48	1.964	-0.0166
OCDM	27.18	7.002	0.7798	25.32	6.029	0.5347	30.26	5.457	0.1640
ΛCDM	26.79	6.173	0.4913	23.48	4.974	0.3122	26.12	4.248	0.0823
ΛCDM2	29.84	7.441	0.7465	26.51	6.093	0.4840	30.23	5.274	0.1393

Note.—Coefficients for polynomial fits (eq. [24]) to the predicted relations between quasar lifetime and clustering length shown in Fig. 7. Col. (1) lists the model name; cols. (2)–(4) the coefficients for $z = 2$; cols. (5)–(7) the coefficients for $z = 3$; and cols. (8)–(10) the coefficients for $z = 4$.

because of their lower values of σ_8 and $D(z)$. The $\Omega_M = 1$ models have larger bias factors, but these are not enough to compensate for the smaller mass fluctuations. The r_1 versus t_Q relations are also shallower for the $\Omega_M = 1$ models, because the values of M_{min} lie farther out on the steep, high-mass tail of the mass function, where a smaller change in M_{min} can make up for the same change in t_Q. The three low-density models yield very similar predictions.

To facilitate comparison of future observational results to our predictions, we have fitted polynomials of the form

$$r_1 = \begin{cases} a_0 + a_1 \log_{10} t_Q & \log_{10} t_Q \geq -1.5, \\ a_0 + a_1 \log_{10} t_Q + a_2(\log_{10} t_Q + 1.5)^2 & \log_{10} t_Q < -1.5, \end{cases}$$

$$(24)$$

to each of the r_1 versus t_Q curves shown in Figure 7. The values of the coefficients are given in Table 3, and the coefficients a_0 and a_1 have the same value over the entire range in t_Q. These fits are accurate to better than 3% in r_1 for given t_Q, or to better than 10% in t_Q given r_1, for all cases except the SCDM and τCDM models at $z = 4$, where the maximum errors are 3% in r_1 given t_Q and 20% in t_Q for a given r_1.

4. DISCUSSION

4.1. Sensitivity to Model Details

As already mentioned in § 2.3, the definition of a "halo lifetime" is somewhat ambiguous. We have so far adopted a definition of t_H as the median time before a halo of mass M is incorporated into a halo of mass $2M$. If we increase this mass ratio from 2 to 5 (a rather extreme value), then the typical halo lifetimes in our CDM models increase by factors of 2–4. Since it is the ratio t_Q/t_H that enters our determination of M_{min} (eq. [8]), and hence fixes the bias factor, this change in t_H would require an equal increase in t_Q to maintain the same clustering length r_1. We conclude that the ambiguity in halo lifetime definition introduces a factor of ~ 2 uncertainty in the determination of t_Q from clustering measurements, in the context of our model.

We have also assumed that quasar luminosity is perfectly correlated with host halo mass, so that matching the space density of an absolute-magnitude–limited sample imposes a sharp cutoff in the host mass distribution at $M = M_{min}$. If there is some scatter in the luminosity–host mass relation, then some halos with $M < M_{min}$ will host a quasar above the absolute-magnitude limit and some halos with $M > M_{min}$ will not. We can model such an effect by introducing a soft cutoff into equation (8):

$$\Phi(z) = \int_0^\infty dM g(M) \frac{t_Q}{t_H(M, z)} n(M, z),$$

$$(25)$$

where

$$g(M) = \begin{cases} 0 & \text{for } M < \dfrac{M_{min}}{\alpha}, \\ \left(\dfrac{\alpha}{M_{min}(\alpha^2 - 1)}\right)M - \dfrac{1}{\alpha^2 - 1} & \text{for } \dfrac{M_{min}}{\alpha} < M < \alpha M_{min}, \\ 1 & \text{for } M > \alpha M_{min}, \end{cases}$$

$$(26)$$

and $\alpha > 1$. Adopting a soft cutoff slightly decreases M_{min} and, more significantly, reduces the value of b_{eff} by allowing some quasars to reside in lower mass halos, which are less strongly biased. Quantitatively, we find that setting $\alpha = 2$, which corresponds to including halos down to $M = M_{min}/2$, decreases the clustering length by $\lesssim 6\%$ for the shortest quasar lifetimes and $\lesssim 10\%$ for the longest quasar lifetimes. Matching a fixed r_1 with an $\alpha = 2$ cutoff requires lifetimes that are longer by a factor of ~ 1–1.5 at short t_Q and ~ 2–2.5 at long t_Q. Longer lifetimes are more sensitive to scatter in the luminosity–host mass relation because b_{eff} depends more strongly on M_{min}/M_* for these rarer objects. The assumption of a perfectly monotonic relation between quasar luminosity and host mass leads to the smallest t_Q for a given r_1. Thus, if any scatter does exist in this relation, our model predictions for t_Q effectively become lower limits to the quasar lifetime.

Another simplification of our model is the assumption that a quasar is either "on" or "off"; each quasar shines at luminosity L for time t_Q, perhaps divided among several episodes of activity, and the rest of the time it is too faint to appear in a luminous quasar sample. More realistically, variations in the accretion rate and radiative efficiency will cause the quasar luminosity to vary, especially if the black hole mass itself grows significantly during the luminous phase. Nonetheless, the maximum luminosity will still depend on the maximum black hole mass. At a given time, the luminous quasar population will include black holes shining at close to their maximum luminosity and "faded" black holes of higher mass. Because the host halos lie on the steeply falling tail of the mass function, the first component of the population always dominates over the second, and we therefore expect our clustering method to yield the time t_Q for which a quasar shines within a factor ~ 2 of its peak luminosity. More strongly faded quasars are too rare to make much difference to the space density or effective bias.

To illustrate this point, we consider the model of Haehnelt et al. (1998) in which a quasar hosted by a halo of mass M has a luminosity history $L(t) = L_0(M) \exp(-t/t_Q)$, with a maximum luminosity $L_0(M) = \alpha M$ proportional to the halo mass. In this model, the time that a quasar shines above the luminosity threshold, $L_{min} = L_0(M_{min})$, of a survey is the visibility time, $t'_Q = t_Q \ln(M/M_{min})$. We can calculate M_{min} for a given space density by substituting t'_Q for t_Q in equation (8), then calculate $b_{eff}(M_{min})$ by multiplying the integrands in the numerator and denominator of equation (13) by the visibility weighting factor $\ln(M/M_{min})$. The middle curves in Figure 8 compare $r_1(t_Q)$ for the on-off (*solid line*) and exponential decay (*dotted line*) models, in the case of ΛCDM at $z = 3$ with our standard $\Phi(z)$. The curves are remarkably similar, showing that the lifetime inferred from clustering assuming an on-off model would be close to the e-folding timescale in an exponential decay model. The curves for the exponential decay model are slightly shallower because at low M_{min} (low t_Q) the mass function is not as steep, allowing faded quasars in more massive halos to make a larger contribution to b_{eff} and thereby raise r_1. Although results for a different functional form of $L(M, t)$ would differ in detail, we would expect the lifetime inferred from clustering to be close to the "half-maximum" width of the typical luminosity history, for the general reasons discussed above.

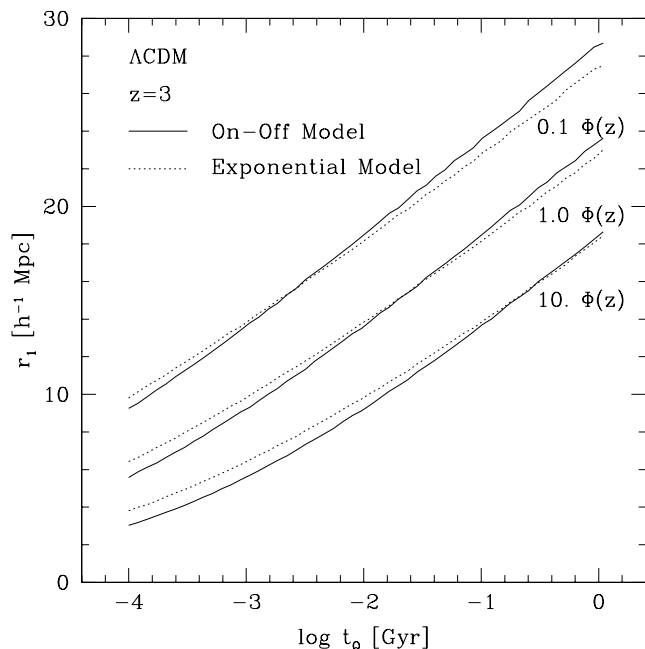As mentioned in § 2.1, we assume that quasars radiate isotropically. If they radiate instead with an average

FIG. 8.—Clustering length vs. t_Q for the ΛCDM model at $z = 3$ for two different models of quasar luminosity evolution at three different space densities $\Phi(z)$. The "on-off" model (solid lines) assumes that the quasar luminosity is constant throughout its lifetime t_Q and is the standard model we discuss in this paper. The central line shows results for the ΛCDM model at $z = 3$ with our standard $\Phi(z)$. The other solid curves, related to the first by simple horizontal shifts, show results for space densities different by factors of 10 and 1/10 (bottom and top, respectively). In the exponential model (dotted lines), the quasar luminosity starts at some maximum luminosity proportional to the halo mass and decays with an e-folding timescale t_Q. The middle line again corresponds to our standard $\Phi(z)$, and the other two dotted curves show results for space densities different by factors of 10 and 1/10 (bottom and top, respectively).

beaming factor $f_B < 1$, then the true value of $\Phi(z)$ is larger than the observed value by a factor f_B^{-1}. The implied lifetime for a given r_1 would therefore be larger by a factor f_B^{-1} as well.

4.2. Interpretation of Existing Data

After several attempts (Osmer 1981; Webster 1982), quasar clustering was first detected by Shaver (1984), and later by Shanks et al. (1987) and Iovino & Shaver (1988). However, measurements of quasar clustering are still hampered by small, sparse samples, and even the best studies to date yield detections with only several σ significance. Given the limitations of current data, it is not surprising that different authors reach different conclusions about the strength of clustering and its evolution. Analyzing a combined sample of quasars with $0.3 < z < 2.2$ from the Durham/AAT UVX Survey, the Canada-France-Hawaii Telescope survey, and the Large Bright Quasar Sample, Shanks & Boyle (1994) and Croom & Shanks (1996) find a reasonable fit to the data with an $\Omega_M = 1$ model that has $\xi(r) = (r/r_0)^{-\gamma}$, $\gamma \approx 1.8$, and a constant comoving correlation length $r_0 = 6\ h^{-1}$ Mpc. La Franca, Andreani, & Cristiani (1998) report a higher correlation length, $r_0 = 9.1 \pm 2.0\ h^{-1}$ Mpc, for a $\gamma = 1.8$ power law, in their $1.4 < z < 2.2$ sample.

If we adopt $r_0 \approx 8\ h^{-1}$ Mpc at $z = 2$ and a corresponding $r_1 \approx 11\ h^{-1}$ Mpc, then the implied quasar lifetime is $\sim 10^{7.5}$

yr for the τCDM model and $\sim 10^8$ yr for SCDM. The r_0 values quoted above are for $\Omega_M = 1$, and because quasar pair separations are measured in angle and redshift, they should be increased by a factor of ~ 1.5 in an $\Omega_M = 0.3$, $\Omega_\Lambda = 0.7$ universe and a factor of ~ 1.3 in an $\Omega_M = 0.3$, $\Omega_\Lambda = 0$ universe (roughly the inverse cube roots of the volume ratios listed in Table A1). Adopting $r_1 \approx 16\ h^{-1}$ Mpc implies a lifetime of $t_Q \sim 10^7$–$10^{7.5}$ yr in our low-density models. However, these numbers must be considered highly uncertain because of the limitations of current data and because the space densities of the various observational samples do not necessarily match those assumed in our model predictions.

All these measurements are based mainly on quasars with $z < 2$. At higher redshift, Kundić (1997) and Stephens et al. (1997) have investigated clustering in the Palomar Transit Grism Survey (PTGS; Schneider, Schmidt, & Gunn 1994). Fitting a $\gamma = 1.8$ power law, Stephens et al. (1997) find $r_0 = 17.5 \pm 7.5\ h^{-1}$ Mpc for $z > 2.7$. This high correlation length (inferred from the presence of three close pairs in a sample of 90 quasars) could be a statistical fluke, but in the context of our model it is tempting to see it as a consequence of the high luminosity threshold of the PTGS survey, which might lead it to pick out the most strongly clustered members of the quasar population.

4.3. Prospects

The 2dF (Boyle et al. 2000; Shanks et al. 2000) and Sloan (York et al. 2000) quasar surveys will transform the study of quasar clustering over the next several years, yielding high-precision measurements for a wide range of redshifts. These measurements will allow good determination of the typical quasar lifetime, t_Q, in the context of the model presented here. They will also test the key assumption of this model, the monotonic relation between quasar luminosity and host halo mass, by characterizing the clustering as a function of redshift and, especially, as a function of quasar absolute magnitude.

Figure 8 illustrates this test for the ΛCDM model at $z = 3$. Brighter quasars have a lower space density $\Phi(z)$, so they should have a higher minimum host halo mass $M_{\min}$, and, because of the higher bias of more massive halos, they should exhibit stronger clustering. Fainter, more numerous quasars should exhibit weaker clustering. Figure 8 shows the predicted r_1 versus t_Q relation for samples with 1/10 and 10 times the space density of our standard case (3.42 quasars per deg^2 per unit redshift; see Table A1). In our standard on-off model (solid lines), a change in $\Phi(z)$ in equation (8) can be exactly compensated for by changing t_Q by the same factor, so the solid curves in Figure 8 are simply shifted horizontally relative to each other. Our predictions in Figure 7 (see eq. [24]) can therefore be transformed to any quasar space density by changing t_Q in proportion to $\Phi(z)$. In the exponential decay model (dotted lines), the scaling of t_Q with $\Phi(z)$ is no longer exact, although it is still a good approximation.

If there is a large dispersion in the relation between quasar luminosity and host halo mass, then the dependence of clustering strength on quasar space density will be much weaker than Figure 8 predicts. Detection of the predicted trend between luminosity and clustering, or a definitive demonstration of its absence, would itself provide an important insight into the nature of quasar host halos. More generally, the parameters of a model that incorporates

scatter (such as the α prescription of eq. [26]) could be determined by matching the observed relation between r_1 and $\Phi(z)$.

If the observations do support a tight correlation between luminosity and host halo mass, then the first property of quasars to emerge from the 2dF and Sloan clustering studies will be the typical lifetime t_Q. For the low-density models in Figure 7, the slope of the correlation between r_1 and $\log_{10} t_Q$ is ~ 10, so a determination of r_1 with a precision of 2 h^{-1} Mpc would constrain t_Q to a factor of $10^{0.2} \approx 1.6$, for a specified cosmology. By the time these quasar surveys are complete, a variety of observations may have constrained cosmological parameters to the point that they contribute negligible uncertainty to this constraint. Instead, the uncertainty in t_Q will probably be dominated by the limitations of the quasar population model, e.g., the approximate nature of the assumptions that the quasar luminosity tracks the halo mass, that there is only one quasar per halo, and that the average lifetime t_Q is independent of quasar luminosity. These assumptions can be tested empirically to some degree, but not perfectly. Despite these limitations, it seems realistic to hope that t_Q can be constrained to a factor of 3 or better by high-precision clustering measurements, a vast improvement over the current situation. It is worth reiterating that our assumption of a perfectly monotonic relation between luminosity and halo mass leads to the smallest t_Q for an observed r_1, since with a shorter lifetime there are simply not enough massive, highly biased halos to host the quasar population.

A determination of t_Q to a factor of 3 will be sufficient to address fundamental issues regarding the physics of quasars and galactic nuclei. Comparison of t_Q to the Salpeter timescale will answer one of the most basic questions about supermassive black holes: do they shine as they grow? If $t_Q \gtrsim 4 \times 10^7$ yr, the e-folding timescale for $L \sim L_E$, $\epsilon \sim 0.1$, then quasar black holes increase their mass by a substantial factor during their optically bright phase. If t_Q is much shorter than this, then the black holes must accrete most of their mass at low efficiency, or while shining at $L \ll L_E$. A short lifetime could indicate an important role for advection-dominated accretion (Narayan, Mahadevan, & Quatert 1998 and references therein), or it could indicate that black holes acquire much of their mass through mergers with other black holes, emitting binding energy in the form of gravitational waves rather than electromagnetic waves. A determination of t_Q would also resolve the question of whether the black holes in the nuclei of local galaxies are the remnants of dead quasars. For example, Richstone et al. (1998) infer a lifetime of $t_Q \sim 10^6$ yr by matching the space density of local spheroids that host black holes of mass $M \gtrsim 4 \times 10^8 \, M_\odot$ to the space density of high-redshift quasars of luminosity $L_E(M) \gtrsim 6 \times 10^{46}$ ergs s^{-1}. If clustering implies a much longer lifetime, then these numerous local black holes may once have powered active nuclei, but they were not the engines of the luminous, rare quasars.

We have assumed in our model that quasar activity is a random event in the life of the parent halo. Quasar activity might instead be triggered by a major merger, by a weaker "fly-by" interaction, or by the first burst of star formation in the host galaxy. Regardless of the trigger mechanism, the lifetime will be the dominant factor in determining the strength of high-redshift quasar clustering, if our assumed link between luminosity and halo mass holds. However, different triggering mechanisms might be diagnosed by more subtle clustering properties, such as features in the correlation function at small separations, or higher order correlations. At low redshift, where the evolution of the quasar population is driven by fueling rather than by black hole growth, the nature of the triggering mechanism might play a major role in determining quasars' clustering properties. The calculations presented here illustrate the promise of quasar clustering as a tool for testing ideas about quasar physics, a promise that should be fulfilled by the large quasar surveys now under way.

APPENDIX A

CONVERTING FROM OBSERVED QUASAR NUMBERS TO $\Phi(z)$

The observed quantity that is measured in studies of quasar clustering and the quasar space density is the number of sources brighter than a given apparent magnitude m per unit redshift per unit solid angle on the sky. This surface density per unit redshift can be converted into a comoving space density of objects brighter than a given absolute magnitude M,

$$\Phi(z, < M) = \frac{dN(<m)}{d\Omega \, dz} \frac{d\Omega \, dz}{dV_c(z)} , \tag{A1}$$

where $dV_c(z)$ is the differential comoving volume element corresponding to $d\Omega \, dz$. Following the notation in Hogg (1999), this volume element is

$$dV_c(z) = D_H \frac{D_M^2}{E(z)} d\Omega \, dz , \tag{A2}$$

TABLE A1
QUASAR SPACE AND SURFACE DENSITY

z (1)	$\Phi(z)$ (2)	$\dfrac{dN}{dz\,d\Omega}$ (3)	$\dfrac{f(1.0, 0.0)}{f(0.3, 0.0)}$ (4)	$\dfrac{f(1.0, 0.0)}{f(0.3, 0.7)}$ (5)
2..............	1.889×10^{-6}	2.132	0.425	0.279
3..............	3.331×10^{-6}	3.417	0.339	0.253
4..............	3.200×10^{-7}	0.287	0.287	0.241

NOTE.—Adopted values of the space density of quasars at $z = 2$, 3, and 4, and cosmological conversion factors. Values of $\Phi(z)$ in col. (2) are from Warren et al. 1994 (assuming $\Omega_M = 1$) for quasars with $M_c < -24.5$ (absolute continuum flux at 1216 Å), converted from their adopted $h = 0.75$ to h^3 Mpc^{-3}. Col. (3) lists the abundance of quasars in number per deg^2 per unit redshift to which the space density in col. (2) corresponds. Cols. (4) and (5) contain the ratios of the factors $f(\Omega_M, \Omega_\Lambda)$ defined in eq. (A6) needed to convert the space density in col. (2), which is valid for $\Omega_M = 1$, to space densities for the OCDM and ΛCDM models, respectively.

where $D_H = c/H_0$ is the Hubble distance, D_M is the transverse comoving distance,

$$D_M = \begin{cases} D_H \dfrac{1}{\sqrt{\Omega_k}} \sinh\left(\sqrt{\Omega_k}\,\dfrac{D_c}{D_H}\right) & \text{for } \Omega_k > 0\,, \\[2ex] D_H \displaystyle\int_0^z \dfrac{dz'}{E(z')} & \text{for } \Omega_k = 0\,, \\[2ex] D_H \dfrac{1}{\sqrt{|\Omega_k|}} \sin\left(\sqrt{|\Omega_k|}\,\dfrac{D_c}{D_H}\right) & \text{for } \Omega_k < 0\,, \end{cases} \tag{A3}$$

and $E(z) = [\Omega_M(1 + z)^3 + \Omega_k(1 + z)^2 + \Omega_\Lambda]^{1/2}$, where $\Omega_k = 1 - \Omega_M - \Omega_\Lambda$. For $\Omega_M = 1$ and $\Omega_k = 0$, the differential comoving volume element is

$$dV_C(z) = 4\left(\dfrac{c}{H_0}\right)^3 (1 + z)^{-3/2} \left(1 - \dfrac{1}{\sqrt{1 + z}}\right)^2 d\Omega\,dz \tag{A4}$$

per steradian per unit redshift.

The fact that $dV_c(z)$ depends on the cosmological parameters means that a given measured surface density of sources corresponds to a different comoving space density for different cosmological model parameters. The space density of quasars is commonly quoted for an $\Omega_M = 1$ universe. To convert this space density (in units of h^3 Mpc^{-3}) into the space density for a model with different values of Ω_M and Ω_Λ requires a correction of the form

$$\Phi(z, \Omega_M, \Omega_\Lambda) = \Phi(z, \Omega'_M, \Omega'_\Lambda) \dfrac{f(z, \Omega'_M, \Omega'_\Lambda)}{f(z, \Omega_M, \Omega_\Lambda)}, \tag{A5}$$

where

$$f(z, \Omega_M, \Omega_\Lambda) = \dfrac{D_H D_M^2}{E(z)}\,. \tag{A6}$$

This procedure converts the reported space density under one set of cosmological parameters back into the observed surface density and then converts the surface density into the space density for the new set of cosmological parameters. In the notation of Popowski et al. (1998), $f(z, \Omega_M, \Omega_\Lambda) = gf^2$, where f and g are given by their equations (5) and (6), respectively.

In Table A1 we list the factors to convert the space density in column (2), which is listed for $\Omega_M = 1$, $\Omega_\Lambda = 0$, to the corresponding space densities for $\Omega_M = 0.3$, $\Omega_\Lambda = 0.0$ and $\Omega_M = 0.3$, $\Omega_\Lambda = 0.7$. The factors in Table A1 are all less than unity because the comoving volume element is smallest in an $\Omega_M = 1$ universe.

REFERENCES

Adelberger, K. L., Steidel, C. C., Giavalisco, M., Dickinson, M., Pettini, M., & Kellogg, M. 1998, ApJ, 505, 18
Bahcall, N. A., Ostriker, J. P., Perlmutter, S., & Steinhardt, P. J. 1999, Science, 284, 1481
Bajtlik, S., Duncan, R. C., & Ostriker, J. P. 1988, ApJ, 327, 570
Bardeen, J., Bond, J. R., Kaiser, N., & Szalay, A. S. 1986, ApJ, 304, 15
Baugh, C. M., & Efstathiou, G. 1993, MNRAS, 265, 145
Bechtold, J. 1994, ApJS, 91, 1

Bond, J. R., Cole, S., Efstathiou, G., & Kaiser, N. 1991, ApJ, 379, 440
Boyle, B. J., Fong, R., Shanks, T., & Peterson, B. A. 1990, MNRAS, 243, 1
Boyle, B. J., Shanks, T., Croom, S. M., Smith, R. J., Miller, L., Loaring, N., & Heymans, C. 2000, MNRAS, 317, 1014
Boyle, B. J., Smith, R. J., Shanks, T., Croom, S. M., & Miller, L. 1999, in IAU Symp. 183, Cosmological Parameters and the Evolution of the Universe, ed. K. Sato (Dordrecht: Kluwer), 178
Bunn, E. F., & White, M. 1997, ApJ, 480, 6

Carroll, S. M., Press, W. H., & Turner, E. L. 1992, ARA&A, 30, 499
Ciotti, L., & Ostriker, J. P. 1999, ApJ, submitted (preprint astro-ph/9912064)
Croft, R. A. C., Weinberg, D. H., Pettini, M., Katz, N., & Hernquist, L. 1999, ApJ, 520, 1
Croom, S. M., & Shanks, T. 1996, MNRAS, 281, 893
Efstathiou, G., & Rees, M. J. 1988, MNRAS, 230, 5P
Eke, V. R., Cole, S., & Frenk, C. S. 1996, MNRAS, 282, 263
Faber, S. M., & Jackson, R. E. 1976, ApJ, 204, 668
Fan, X., et al. 1999, AJ, 118, 1
Haehnelt, M. G., Natarajan, P., & Rees, M. J. 1998, MNRAS, 300, 817
Haehnelt, M. G., & Rees, M. J. 1993, MNRAS, 263, 168
Haiman, Z., & Hui, L. 2000, ApJ, in press (preprint astro-ph/0002190)
Haiman, Z., & Loeb, A. 1998, ApJ, 503, 505
Heath, D. J. 1977, MNRAS, 179, 351
Hewett, P. C., Foltz, C. B., & Chaffee, F. C. 1993, ApJ, 406, L43
Hogg, D. W. 1999, preprint (astro-ph/9905116)
Iovino, A., & Shaver, P. A. 1988, ApJ, 330, L13
Jing, Y. P. 1998, ApJ, 503, L9
Kaiser, N. 1984, ApJ, 284, L9
Katz, N., Hernquist, L., & Weinberg, D. H. 1999, ApJ, 523, 463
Katz, N., Quinn, T., Bertschinger, E., & Gelb, J. M. 1994, MNRAS, 270, L71
Kauffmann, G., & Haehnelt, M. 2000, MNRAS, 311, 576
Kolatt, T. S., et al. 1999, ApJ, 523, L109
Kundić, T. 1997, ApJ, 482, 631
Lacey, C., & Cole, S. 1993, MNRAS, 262, 627
La Franca, F., Andreani, P., & Cristiani, S. 1998, ApJ, 497, 529
Lynden-Bell, D. 1969, Nature, 223, 690
Magorrian, J., et al. 1998, AJ, 115, 2285
McDonald, P., Miralda-Escudé, J., Rauch, M., Sargent, W. L. W., Barlow, T. A., Cen, R., & Ostriker, J. P. 2000, ApJ, 543, 1
Melchiorri, A., et al. 2000, ApJ, 536, L63
Miller, A. D., et al. 1999, ApJ, 524, L1
Mo, H. J., Mao, S., & White, S. D. M. 1999, MNRAS, 304, 175
Mo, H. J., & White, S. D. M. 1996, MNRAS, 282, 347 (M-W)
Monaco, P., Salucci, P., & Danese, L. 2000, MNRAS, 311, 279
Narayan, R., Mahadevan, R., & Quatert, E. 1998, in The Theory of Black Hole Accretion Discs, ed. M. A. Abramowicz, G. Bjornsson, & J. E. Pringle (Cambridge: Cambridge Univ. Press), 148

Navarro, J. F., Frenk, C. S., & White, S. D. M. 1997, ApJ, 490, 493
Osmer, P. S. 1981, ApJ, 247, 762
———. 1998, in ASP Conf. Ser. 146, The Young Universe, ed. S. D'odorico, A. Fontana, & E. Giallongo (San Francisco: ASP), 1
Peacock, J. A., & Dodds, S. J. 1994, MNRAS, 267, 1020
Peebles, P. J. E. 1980, The Large Scale Structure of the Universe (Princeton: Princeton Univ. Press)
Perlmutter, S., et al. 1999, ApJ, 517, 565
Phillips, J., Croft, R. A. C., Weinberg, D. H., Hernquist, L., Katz, N., & Pettini, M. 2000, ApJ, submitted (preprint astro-ph/0001089)
Press, W. H., & Schechter, P. 1974, ApJ, 187, 425 (P-S)
Popowski, P. A., Weinberg, D. H., Ryden, B. S., & Osmer, P. S. 1998, ApJ, 498, 11
Richstone, D., et al. 1998, Nature, 395, A14
Riess, A. G., et al. 1998, AJ, 116, 1009
Salpeter, E. E. 1964, ApJ, 140, 796
Salucci, P., Szuszkiewicz, E., Monaco, P., & Danese, L. 1999, MNRAS, 307, 637
Schneider, D. P., Schmidt, M., & Gunn, J. E. 1994, AJ, 107, 1245
Shanks, T., & Boyle, B. J. 1994, MNRAS, 271, 753
Shanks, T., Boyle, B. J., Croom, S. M., Loaring, N., Miller, L., & Smith, R. J. 2000, in ASP Conf. Ser. 200, Clustering at High Redshift, ed. A. Mazure, O. LeFevre, & V. Lebrun (San Francisco: ASP), in press
Shanks, T., Fong, R., Boyle, B. J., & Peterson, B. A. 1987, MNRAS, 227, 739
Shaver, P. A. 1984, A&A, 136, L9
Stephens, A. W., Schneider, D. P., Schmidt, M., Gunn, J. E., & Weinberg, D. H. 1997, AJ, 114, 41
Sugiyama, N. 1995, ApJS, 100, 281
Tegmark, M., & Zaldarriaga, M. 2000, ApJ, 544, 30
Turner, E. L. 1991, AJ, 101, 5
van der Marel, R. P. 1999, AJ, 117, 744
Warren, S. J., Hewett, P. C., & Osmer, P. S. 1994, ApJ, 421, 412
Webster, A. 1982, MNRAS, 199, 683
Weinberg, D. H., Croft, R. A. C., Hernquist, L., Katz, N., & Pettini, M. 1999, ApJ, 522, 563
White, S. D. M., Efstathiou, G. P., & Frenk, C. S. 1993, MNRAS, 262, 1023
York, D., et al. 2000, AJ, in press
Zeldovich, Ya. B., & Novikov, I. D. 1964, Soviet Phys. Dokl., 158, 811