SELECTION BIAS IN OBSERVING THE COSMOLOGICAL EVOLUTION OF THE $M_\bullet-\sigma$ AND $M_\bullet-L$ RELATIONSHIPS

TOD R. LAUER
National Optical Astronomy Observatory,1 P.O. Box 26732, Tucson, AZ 85726

SCOTT TREMAME
Institute for Advanced Study, Einstein Drive, Princeton, NJ 08540

DOUGLAS RICHSTONE
Department of Astronomy, University of Michigan, Ann Arbor, MI 48109

AND

S. M. FABER
UCO/Lick Observatory, Board of Studies in Astronomy and Astrophysics, University of California, Santa Cruz, CA 95064

Received 2007 May 28; accepted 2007 August 1

ABSTRACT

Programs to observe evolution in the $M_\bullet-\sigma$ or $M_\bullet-L$ relations typically compare black hole masses, $M_\bullet$, in high-redshift galaxies selected by nuclear activity to $M_\bullet$ in local galaxies selected by luminosity $L$ or velocity dispersion $\sigma$. Because active galactic nucleus luminosity can depend on $M_\bullet$, selection effects are different for the two samples, potentially producing a false signal of evolution. Cosmic scatter in the $M_\bullet-\sigma$ relations means that the mean log $L$ or log $\sigma$ among galaxies that host a black hole of given $M_\bullet$ may be substantially different than the log $L$ or log $\sigma$ obtained from inverting the $M_\bullet-L$ or $M_\bullet-\sigma$ relations for the same nominal $M_\bullet$. The bias is strongest at high $M_\bullet$, where the luminosity and dispersion functions of galaxies fall rapidly. The most massive black holes occur more often as rare outliers in galaxies of modest mass than in the even rarer high-mass galaxies, which would otherwise be the sole location of such black holes in the absence of cosmic scatter. Because of this bias, $M_\bullet$ will typically appear to be too large in the distant sample for a given $L$ or $\sigma$. For the largest black holes and the largest plausible cosmic scatter, the bias can reach a factor of 3 in $M_\bullet$ for the $M_\bullet-\sigma$ relation and 9 for the $M_\bullet-L$ relation. Unfortunately, the cosmic scatter is not known well enough to correct for the bias. Measuring evolution of the $M_\bullet$ relations requires object selection to be precisely defined and exactly the same at all redshifts.

Subject headings: galaxies: evolution — galaxies: fundamental parameters — galaxies: nuclei

1. OBSERVING EVOLUTION IN THE RELATIONSHIPS BETWEEN BLACK HOLE MASS AND GALAXY PROPERTIES

The discovery that most elliptical galaxies and spiral bulges host a black hole at their centers, plus the tight relations observed between black hole mass $M_\bullet$ and galaxy luminosity $L$ or stellar velocity dispersion $\sigma$ (Dressler 1989; Kormendy 1993; Kormendy & Richstone 1995; Magorrian et al. 1998; Ferrarese & Merritt 2000; Gebhardt et al. 2000a; Tremaine et al. 2002; Häring & Rix 2004), suggests that the formation and growth of central black holes is deeply intertwined with that of their host galaxies. Recent theoretical work (e.g., Hopkins et al. 2006) supports this view, and also predicts how the relations between the properties of black holes and their host galaxies have changed over time. Direct observation of the evolution of the $M_\bullet-\sigma$ and $M_\bullet-L$ relations over cosmological time would offer unique insight into galaxy and black hole formation.

There have been many attempts to measure the evolution of the $M_\bullet$ relations; we cite a partial list of these:

1. Shields et al. (2003) examined a sample of quasars at redshifts up to 3.3. The luminosity $L$ and velocity dispersion $\sigma$ of the host galaxy in such objects cannot be measured reliably because the light from the galaxy is overwhelmed by the quasar flux; instead, they use the width of the narrow [O iii] emission line as a surrogate for $\sigma$. They estimate the black hole mass using the photoionization method, which is based on an empirically calibrated relation involving the continuum luminosity of the active galactic nucleus (AGN) and the width of the H\beta or other broad emission lines.2 The theoretical assumptions that motivated this relation are that the bulk velocities of the emitting clouds are determined by their orbital motion in the gravitational field of the black hole, and that the emitting region is photoionized by an ultraviolet continuum spectrum of fixed shape. In practice, however, it is simply an empirical relation defined and calibrated by reverberation mapping of objects at very low redshift. They find an $M_\bullet-\sigma$ relation that is consistent with the local one, suggesting that this relation is independent of redshift.

2. Using similar methods on a larger but lower redshift quasar sample from the Sloan Digital Sky Survey (SDSS), Salvatier et al. (2007) found that galaxies of a given dispersion at $z \approx 1$ have black hole masses that are larger by $\Delta \log M_\bullet \sim 0.2$ than at $z = 0$.

3. Treu et al. (2004) and Woo et al. (2006) measured the velocity dispersion of 14 Seyfert 1 galaxies at redshift $z \approx 0.36$;

2 Historically this method was based on photoionization calculations. Shields et al. (2003) and other authors, however, used a purely empirical method that assumes a particular form for the broad-line region radius-luminosity relation, calibrated by reverberation observations. It might be more appropriate to call this an “empirical single-epoch method.”
they used stellar absorption lines, which should provide a more
direct measure of the dispersion than emission lines. They mea-
sured the black hole mass using the photoionization method.
They found that galaxies of a given dispersion at \(z \simeq 0.36\) host
black holes having \(\Delta \log M_\bullet = 0.62 \pm 0.10 \pm 0.25\).

4. Peng et al. (2006a) and Peng et al. (2006b) measured the
bulge luminosities of a combined sample 51 AGN hosts at \(z > 1\)
using the Hubble Space Telescope (HST), estimating their black
hole masses using the photoionization method. They concluded
that the \(M_\bullet - L\) relation at \(z \sim 2\) is close to the relation at \(z = 0\),
finding \(\Delta \log M_\bullet \lesssim 0.3\). This result is remarkable since black
holes can only grow with time, while elliptical galaxies fade with
time as their stars die: even if the black holes do not grow at all,
passive stellar evolution models would predict \(\Delta \log M_\bullet\) between
\(-0.4\) and \(-0.8\) at \(z \approx 2\). If their result is correct, then either black
holes must be ejected from the galaxy centers and replaced with
smaller ones or the galaxy luminosity must grow substantially
through mergers.

A common thread among all these investigations and others is
that the high-redshift sample is selected by some measure of AGN
visibility. Black hole masses and galaxy properties are then de-
nered and an \(M_\bullet - \sigma\) or \(M_\bullet - L\) relation is fitted to the data. Since
the low-redshift relations are linear in log \(M_\bullet\), log \(L\), and log \(\sigma\)
( eqs. [4] or [5] below), the high-redshift data are often analyzed by
assuming that the slope of the relation is the same as at low
redshift and estimating the offset in the intercept or zero point,
which may be expressed as \(\Delta \log M_\bullet\) at fixed \(L\) or \(\sigma\). Any such
offset is then interpreted as evidence that the ratio of \(M_\bullet\) to \(\sigma\) or
\(L\) has changed over time. Of course, the offset can be (and some-
times is) equally well expressed as \(\Delta \log L\) or \(\Delta \log \sigma\) at fixed
black hole mass.

Proof of evolution in the \(M_\bullet\) relations requires demonstrating that
the high- and low-redshift galaxy samples were assem-
bled with no selection effects that would bias the relation be-
 tween the typical galaxy properties and black hole masses being
measured—or, at least, that the selection effects in the high- and
low-redshift samples were the same. This may be a more difficult
task than has commonly been assumed. The local black hole sam-
ple has mostly been drawn from “normal” galaxies with quies-
cent or low levels of nuclear activity. Despite the proximity of
these galaxies, detecting and weighing their central black holes
requires exquisite spatial resolution, very high signal-to-noise
ratio observations, and elaborate modeling of the observations.
Investigating the properties of black holes in nearby galaxies re-
 mains a frontier problem; to date, dynamically determined black
hole masses are available for only slightly more than three dozen
galaxies. Since there was little evidence of a black hole in most
of these galaxies before HST observations were taken (except
sometimes for a modest rise in the velocity dispersion in spectra
taken at ground-based resolution), and since black holes are
found in almost all nearby galaxies that have been examined care-
fully with HST, the black holes in local quiescent galaxies are
selected mainly on the basis of galaxy properties such as \(L\) or \(\sigma\).

The techniques used to measure black hole masses in quiescent
galaxies at low redshift cannot be used at high redshift, both be-
cause the radius of influence of the black hole cannot be resolved
beyond a few tens of megaparsecs and because the galaxies have
substantially lower surface brightness, by the factor \((1 + z)^3\)2. Black
holes at high redshift are identified instead by their association
with AGNs. The observer thus first locates a black hole that is
accreting matter and weighs it by the properties of the AGN
emission lines, and then second attempts the (still difficult) task
of measuring the properties of the host galaxy. The existence and
properties of the AGN depend both on the properties of the black
hole (mass, spin, and orientation) and on the properties of the
galaxy (mass inflow rate, orientation, etc.). In short, selection is
done at low redshift by galaxy properties and at high redshift by
a combination of black hole and galaxy properties that depends
on the method used.

The determination of black hole masses from the AGN
emission-line spectrum is a difficult task but one that we do not
examine in detail here. The methods used are believed to have
uncertainties of a factor of 4 or so (Vestergaard 2006), but (1) the
physical basis for these methods is not well understood; (2) the
error distribution and its possible dependence on properties such
as orientation or AGN luminosity are poorly known; (3) the masses
are calibrated using reverberation mapping, which in turn is cali-
brated using kinematic measurements of black hole masses, and
each of these methods has its own substantial systematic and
random uncertainties.

Galaxy samples obtained with different selection techniques
will generally satisfy different \(M_\bullet\) relations. This effect, analogous
to the Malquist (1924) bias that is familiar in studies of Galactic
structure, arises because of the cosmic scatter in the \(M_\bullet\) relations;
there is not (so far as we know) an exact one-to-one relation be-
tween black hole mass and any single measurable property of
galaxies such as \(L\) or \(\sigma\). In this paper we argue that determina-
tions of the redshift evolution of the \(M_\bullet\) relations may be strongly
biased by selection effects unless the same sample selection tech-
niques are used at all redshifts. Correcting for the selection bias
introduced by the use of different sample selection criteria at dif-
ferent redshifts is extremely difficult: to make accurate corrections
it is necessary to know both how the selection depends on the gal-
axy and black hole properties and the cosmic scatter in the \(M_\bullet\)
relations. At present we have only crude upper limits to the latter
quantity. We show that even if the cosmic scatter were known, the
bias in the samples described above usually can only be corrected
with additional information, such as how galaxy properties and
black hole mass determine the probability distribution of AGN
luminosity.

The selection bias can be especially large at large \(M_\bullet\), as the
following argument shows. Consider the \(M_\bullet - L\) relation. At high \(L\)
the number density of galaxies falls off rapidly, as illustrated by the
steep cutoff in the Schechter (1976) luminosity function. Cosmic
scatter in the \(M_\bullet - L\) relation implies that there is a distribution
of black hole masses at a given \(L\). The rare high-mass black holes
can arise from either the peak of this distribution in the rare gal-
axies with large \(L\), or from the high-mass tail of the distribution in
the more numerous galaxies of modest \(L\). If the number density
of galaxies is falling off rapidly with \(L\), the contribution from gal-
axies with modest \(L\) may actually overwhelm the population of
black holes of similar mass associated with galaxies of higher \(L\).
This problem was first explicitly identified by Willott et al. (2005)
and Fine et al. (2006) in the context of the correlation between
black hole mass and dark matter halo mass, and by Salviander
et al. (2007) and Treu et al. (2007) in the context of the \(M_\bullet - \sigma\)
relation, but apparently has not been appreciated by most ob-
servers studying the evolution of the \(M_\bullet\) relations.

In this paper we present a more general exposition of the biases
incurred when studying evolution of the \(M_\bullet\) relations by com-
paring samples obtained by different selection criteria and at dif-
ferent redshifts. We start by considering the extraction of samples
from hypothetical joint distributions of \(M_\bullet\) and \(L\) or \(\sigma\). We then
consider the selection bias that occurs if the high-redshift samples
are selected by AGN flux or luminosity. We discuss the prospects
of correcting for selection bias, and argue that selection bias may place fundamental limits on the determination of the evolution of the $M_*$ relations. We conclude with a brief review of attempts to measure the evolution of the $M_*$ relations and how they may have been affected by selection bias.

2. THE JOINT DISTRIBUTION OF BLACK HOLE MASS AND GALAXY PROPERTIES

Understanding object selection bias requires knowledge of the joint probability distribution of black hole mass $M_*$ and a galaxy property $s$, which for the present discussion is either $\log L$, where $L$ is the rest-frame $V$-band galaxy luminosity, or $\log \sigma$, where $\sigma$ is the line-of-sight velocity dispersion in the main body of the galaxy (the precise definition of $L$ or $\sigma$ does not concern us, so long as it is defined consistently in all samples). The true form of the joint distribution unfortunately is unknown; however, we can construct a hypothetical form of this distribution that accurately represents the present state of the observations, and in any case suffices to show how sample selection bias can occur.

We write the probability of finding a galaxy in the interval $(\mu, \mu + d\mu)$ and $(s, s + ds)$ as $\nu(\mu, s) d\mu ds$, where $\mu = \log(M_*)$. From the definition of conditional probability this can be rewritten as

$$\nu(\mu, s) = \nu(\mu | s) g(s),$$

where $g(s) ds$ is the probability that a randomly chosen galaxy in a given volume lies in the interval $(s, s + ds)$ and $\nu(\mu | s) d\mu$ is the probability that the black hole mass lies in the range $(\mu, \mu + d\mu)$ given that the galaxy property is $s$. The function $g(s)$ is then given by either the volume-limited luminosity function—more properly, the luminosity function of early-type galaxy components (ellipticals and spiral bulges), since these are the components that correlate with black hole mass—or the velocity-dispersion distribution of early-type components. Without loss of generality, we may write $\nu(\mu, s) = h(\mu - f(s), s)$, where $\int h(x, y) dx = 1$ and $\int x h(x, s) dx = 0$, so $f(s)$ is the mean value of $\mu$ at a given value of the galaxy property $s$; that is, $f(s)$ is either the $M_*$-L or $M_*$-$\sigma$ relation. The limited observational data on black hole masses in local galaxies is consistent with the hypothesis that $h(x, s)$ is independent of $s$; adopting this hypothesis for simplicity we have

$$\nu(\mu, s) = h[\mu - f(s)] g(s).$$

The intrinsic variance or “cosmic scatter” in black hole mass at a given value of galaxy property $s$ is $\sigma_x^2 = \int x^2 h(x, s) dx$. Observational errors may make $h(x, s)$ appear to be yet broader, but for this discussion, we assume that observational errors are negligible; the selection effects that we are concerned with are solely due to the fact that there is an intrinsic and irreducible range of $M_*$ at any $s$.

It should be stressed that although the functional form in equation (2) is consistent with the available data, it is far from unique. As a foil, we may consider the form

$$\nu(\mu, s) = p(s - c(\mu)) \Phi_*(\mu).$$

A physical model that motivates equation (2) is one in which the galaxy property $s$ determines the black hole mass $\mu$ with a cosmic scatter described by the function $h$, while (3) is motivated by models in which the black hole mass $\mu$ determines the galaxy property $s$, with cosmic scatter described by the function $p$. If $\int p(y) dy = 1$, then $\Phi_*(\mu) d\mu$ is the number of black holes per unit volume with log mass in the range $(\mu, \mu + d\mu)$. We do not know which of equations (2) or (3) is correct (possibly neither, or both); our motivation for choosing the former is that it provides a simple representation of what local observations of black holes in inactive galaxies measure.

Of the three functions that contribute to $\nu(\mu, s)$, $g(s)$ is probably the best determined. We discuss the galaxy-property functions in detail in Lauer et al. (2007) but summarize them briefly here. For the galaxy luminosity function we use the Blanton et al. (2003) SDSS luminosity function, transformed to the $V$-band and redshift $z = 0.1$, augmented with the Postman & Lauer (1995) luminosity function of brightest cluster galaxies, which appear to be undercounted in the SDSS. We note that the Blanton et al. function refers to total rather than bulge luminosity in S0 and spiral galaxies; however, this difference is less important at the bright end of the luminosity function, where the selection effects are strongest. For the velocity-dispersion function we use the Sheth et al. (2003) SDSS results, augmented at the high-$\sigma$ end as prescribed by Bernardi et al. (2006) to account for an artificial high-$\sigma$ cutoff in Sheth et al. Both functions are shown in Figure 1.

The function $f(s)$ should be the least-squares fit of $\log M_*$ to $\log L$ or $\log \sigma$ in a sample selected by galaxy properties. The $M_*$-$\sigma$ relation $f(s)$ is based on the galaxy sample from Tremaine et al. (2002) augmented by a few galaxies with more recent $M_*$ determinations (see Lauer et al. 2007). In contrast to the treatment in Tremaine et al. (2002), which treated $M_*$ and $\sigma$ symmetrically, the appropriate treatment for our purposes is a least-squares fit of $\log M_*$ on $\log \sigma$, which gives

$$\log \left( \frac{M_*}{M_{\odot}} \right) = (4.13 \pm 0.32) \log \left( \frac{\sigma}{200 \text{ km s}^{-1}} \right) + 8.29 \pm 0.07,$$

for $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$ (which we use throughout this paper). We discuss the $M_*$-L relation in detail in Lauer et al. (2007); in brief, we use the Häring & Rix (2004) relation between $M_*$ and galaxy mass, transformed back to luminosity by adopting the mass-to-light ratio $M/L_V \approx 6 \times 10^{-0.092(M_*+22)} M_{\odot}/L_{\odot}$, based on the $M/L$ estimates given in Gebhardt et al. (2003). A least-squares fit of $M_*$ on $M_L$ for galaxies with $M_L < -19$ gives

$$\log \left( \frac{M_*}{M_{\odot}} \right) = (1.32 \pm 0.14) \frac{(M_L - 22)}{2.5} + 8.67 \pm 0.09.$$

Equations (4) and (5) have the form

$$f(s) = a + bs,$$

where $a$ and $b$ are constants. Lauer et al. (2007) showed that the $M_*$-$\sigma$ and $M_*$-L relations make different predictions for $M_*$ in luminous or high-dispersion galaxies, which have lower $\sigma$ values for a given luminosity than is implied by eliminating $M_*$ from equations (4) and (5). Thus at least one of the $M_*$-$\sigma$ or $M_*$-L relations must curve away from the linear fit in equation (6) in the most luminous galaxies. We are not concerned with this issue here and simply show the different selection biases that can result from using $\sigma$ or $L$ as the galaxy property that correlates with $M_*$, assuming that the linear relation (6) holds at all masses. However, the reader should bear in mind that neither the $M_*$-L nor $M_*$-$\sigma$ relation is well determined at high galaxy masses, where the potential selection bias is most important.
The scatter function, $h[\mu - f(s)]$, is poorly known at best. For this analysis, we assume that at any galaxy $L$ or $\sigma$, $\log M_\star$ is described by a normal distribution about $f(s)$, with cosmic scatter $\sigma_\mu$. Thus,

$$h[\mu - f(s)]d\mu = \frac{d\mu}{\sqrt{2\pi}\sigma_\mu^2} \exp\left\{ -\frac{(\mu - f(s))^2}{2\sigma_\mu^2}\right\} \tag{7}$$

Again, we emphasize that $\sigma_\mu$ does not embody any observational errors in the determination of $M_\star$; it represents the intrinsic spread in $M_\star$ at any galaxy property. Novak et al. (2006) concluded that only an upper limit to $\sigma_\mu$ is presently known for either the $M_\star$-$\sigma$ or $M_\star$-$L$ relation, both because of the small sample of reliable $M_\star$ determinations and because of uncertainties in the observational errors in $M_\star$, which must be accurately measured to isolate the contribution of $\sigma_\mu$ to the total residuals. For the present analysis we explore the sensitivity of the selection biases to $\sigma_\mu$ on the assumption that $\sigma_\mu < 0.3$ for dispersion $\sigma$ and $\sigma_\mu < 0.5$ for luminosity $L$.

Apart from the poor knowledge of $\sigma_\mu$, there are at least two other major uncertainties in the function $h$: (1) a more general treatment would allow for the possibility that it varies with $s$, and (2) there is little or no justification of the assumed normal form. As is shown below, for the most massive black holes or galaxies the selection bias may depend on the form that $h[\mu - f(s)]$ takes at several standard deviations away from the mean, yet an observational determination of the form of $h$ in this region would require a sample of well-determined black hole masses several orders of magnitude larger than is presently available. Thus, at the highest black hole masses, the selection bias is likely to depend sensitively on knowledge that is not presently at hand, so it is not possible to apply reliable corrections for this bias.

With these caveats, in Figure 2 we show estimates of $\nu(\mu, s)$ as modeled by equation (2) for each of the $M_\star$-$\sigma$ and $M_\star$-$L$ relations, and two assumed values of $\nu(\mu, s)$. Projection of $\nu(\mu, s)$ onto the $\mu$-axis produces the mass function of black holes,

$$\Phi_\star(\mu) = \int \nu(\mu, s) ds \tag{8}$$

$$= \int h[\mu - f(s)]g(s)ds. \tag{9}$$

The importance of cosmic scatter for converting the luminosity or dispersion function $g(s)$ to a black hole mass function has been discussed several times (Yu & Tremaine 2002; Yu & Lu 2004; Tundo et al. 2007; Lauer et al. 2007). What may be less appreciated, however, is that cosmic scatter implies that the most massive black holes are often hosted by modest galaxies that a priori would not be expected to harbor black holes of high mass. To illustrate this point, Figure 3 shows $\Phi_\star(\mu)$ for three different versions of $\nu(\mu, s)$. The salient feature is that as the cosmic scatter $\sigma_\mu$ increases, the contribution to the density of the most massive black holes from the wings of the scatter function $h$ overwhelms the “native” population of massive black holes harbored by the galaxies with the largest values of property $s$. A second illustration of this point occurs in the top two panels in Figure 4, which show the probability distribution of $L$ and $\sigma$ for $M_\star = 10^{10} M_\odot$; note that the solid curves, corresponding to $\sigma_\mu = 0.5$ for $L$ and $0.3$ for $\sigma$, have a prominent shoulder to the right of the peak, the peak arising from low-luminosity or low-dispersion galaxies, and the shoulder from high luminosities and dispersions.

The source of the selection bias can be seen more directly in Figure 2, which shows $\nu(\mu, s)$ for each of the $M_\star$-$\sigma$ and $M_\star$-$L$ relations, and two assumed values of $\nu(\mu, s)$. Only for $\sigma_\mu = 0$ is there an exact relation between $M_\star$ and $\sigma$ or $L$. Our model for $\nu(\mu, s)$, which may not be correct, but is consistent with the data, implies that the density contours in Figure 2 are symmetric about the mean relationship ridgelines in the vertical direction; in other words, the...
conditional probability of $\mu$ given $s$ is symmetric about $\mu = f(s)$. On the other hand, the conditional probability of $s$ given $\mu$,

$$
\nu(s|\mu) = \frac{\nu(\mu, s)}{\int \nu(\mu, s) \, ds},
$$

is not symmetric about $s = f^{-1}(\mu)$, as seen in Figure 4, which shows the distribution of $s$ at selected values of $M_\bullet$. As $M_\bullet$ and $\sigma_\mu$ increase, the distribution of $s$ for a given $M_\bullet$ moves further and further away from the value implied by the inverse of the $M_\bullet$ relations, $f^{-1}(\mu) = (\mu - a)/b$, where $a$ and $b$ are the coefficients given in equation (6). For $M_\bullet > 10^9 M_\odot$, $f^{-1}(\mu)$ generally falls well out in the wings of the distribution of $s$ for both $L$ and $\sigma$, particularly for the larger values of $\sigma_\mu$. The galaxies with $s = f^{-1}(\mu)$ for the highest-mass black holes are so far down in the step cutoffs of the $L$ and $\sigma$ distribution functions that they are completely overwhelmed by the population of galaxies of modest mass that harbor high-mass black holes as statistical outliers.

Fig. 2.—Joint probability distribution of galaxy luminosity $L$ (top panels) or stellar velocity dispersion $\sigma$ (bottom panels) and black hole mass, $M_\bullet$. The solid, straight, black line gives the mean $M_\bullet$-$L$ or $M_\bullet$-$\sigma$ relation (eqs. [4] and [5]). The joint distribution is calculated by assuming that the black hole mass has a log-normal distribution about the $M_\bullet$-$L$ or $M_\bullet$-$\sigma$ relation, with the normalization provided by the galaxy luminosity or velocity dispersion functions shown in Fig. 1. The adopted dispersions in log $M_\bullet$ are $\sigma_\mu = 0.25, 0.50$ for the $M_\bullet$-$L$ relation or $\sigma_\mu = 0.15, 0.30$ for the $M_\bullet$-$\sigma$ relation. The contours are arbitrary but are spaced in increments of 0.5 dex in density. The curved red lines give the mean $M_\bullet$ or log ($\sigma$) as a function of $M_\bullet$. The displacement of this line from the mean relations gives the bias at a given $M_\bullet$. 

No. 1, 2007 EVOLUTION OF BLACK HOLE MASS RELATIONSHIPS
The mean value of $s$ at a given $M_\bullet$ is

$$\langle s \rangle_\mu = \int s \nu(s|\mu) \, ds = \frac{\int s \nu(s|\mu) \, ds}{\int \nu(s|\mu) \, ds}$$

$$= \frac{\int s \, g(s) h[\mu - f(s)] \, ds}{\int g(s) h[\mu - f(s)] \, ds}.$$  \hspace{1cm} (11)

This is shown as the red lines in Figure 2. The mean of log $L$ or log $\sigma$ at a given $M_\bullet$ is offset from $f^{-1}(\mu)$, which is just the $s$ location of the $M_\bullet$ relation ridgelines shown in the figure. In the presence of cosmic scatter, the mean $s$ of a sample of galaxies that host a black hole of given $\mu$ is different from the mean $\mu$ hosted by a sample of galaxies of given $s$. This difference is the source of the selection bias.

### 3. Illustration of Selection Biases

Selection bias typically occurs because the galaxy samples used to probe the $M_\bullet$ relations at cosmological distances are not selected by galaxy $L$ or $\sigma$, but by the visibility of their AGNs. For most of the discussion in this section we assume that AGN luminosity does not depend directly on $L$, $\sigma$, or any other galaxy property unrelated to $M_\bullet$. This model is appropriate if, for example, the probability that the AGN associated with a black hole of mass $M_\bullet$ has luminosity $L_{AGN}$ is given by

$$dp = \psi(\lambda - \mu) \, d\lambda,$$

where \(\int \psi(\lambda) \, d\lambda = 1,$$ \hspace{1cm} (12)

and $\lambda \equiv \log L_{AGN}$ and $\mu \equiv \log M_\bullet$. The physical content of this assumption is that the luminosity history of an AGN is determined by the black hole mass and scales with the Eddington luminosity, which is proportional to $M_\bullet$.

### 3.1. Bias in a Luminosity-limited Survey

We first consider the bias that occurs in a sample in which (1) all of the objects are in a narrow redshift range, and (2) the survey contains all AGNs brighter than a given flux. In this case the probability that an AGN in the relevant redshift range is accepted in the survey depends only on its luminosity $L_{AGN}$, and by equation (12) this in turn depends only on its black hole mass $\mu$. Thus the probability distribution of a galaxy property $s$ at a given value of $\mu$ is not biased by the selection, and is simply $\nu(s|\mu)$ (eq. [10]). Some examples are shown in Figure 4, which plots the probability distribution of $s - f^{-1}(\mu)$ for several values of black hole mass $M_\bullet$ and cosmic scatter $\sigma_\mu$. Here $f^{-1}(s)$ is the inverse of the $M_\bullet$-L or $M_\bullet$-$\sigma$ relation (eq. [2]). The mean values of the distributions $\langle s \rangle_\mu$ are displaced to values lower than $f^{-1}(\mu)$ in most of the examples shown. As expected, the width and offset of the distributions are larger for larger values of $\sigma_\mu$. Also as expected, the offsets are generally larger for more massive black holes. Figure 2 (red lines) shows $\langle s \rangle_\mu$; the offset of these lines from the relation $\mu = f(s)$ demonstrates the selection bias.

In this model, the selection bias can be corrected: an assumed form for the probability distribution $\nu(\mu, s)$ such as equation (2) with redshift-dependent parameters can be used to compute $\nu(\mu|s)$, which can then be fitted to the distribution of galaxy property $s$ in the sample at a given value of black hole mass $\mu$ [separate issues, discussed in \S 4, are whether this model for the selection effects is realistic and whether the assumed form of $\nu(\mu, s)$ is correct]. However, most papers in the literature have not taken this approach, so it is worthwhile to estimate the biases that might be introduced by using simpler statistics to estimate the evolution in the $M_\bullet$ relations.

Figure 5 shows the object selection bias $\Delta s = f^{-1}(\mu) - \langle s \rangle_\mu$ as a function of $M_\bullet$ and $\sigma_\mu$. In this example we have expressed the bias in terms of galaxy properties, since we have selected...
by $M_\bullet$; however, it may also be represented as $\Delta \log M_\bullet = b \Delta s$, where $b$ is the slope given in equation (6). Note that the bias is not always a monotonic function of black hole mass, nor does it always have the same sign. In this simple model, the bias at black hole masses of $10^9 M_\odot$ is $\Delta \log M_\bullet = 0.4$ in the $M_\bullet$-$\sigma$ relation for $\sigma_\mu = 0.3$, and $\Delta \log M_\bullet = 0.7$ in the $M_\bullet$-L relation for $\sigma_\mu = 0.5$, and even larger at larger black hole masses. The bias is smaller at lower L or $\sigma$, but we caution that the near-zero bias in the $M_\bullet$-L relation at small $M_\bullet$ is an artifact of our assumption that the luminosity function has the Schechter (1976) form at low luminosities (see below), and this may not be true for the early-type galaxy components that are believed to host the black holes.

Figure 5 also shows that a significant portion of the bias comes from galaxies with $|s - f^{-1}(\mu)| > 2\sigma_\mu$ for the larger values of $\sigma_\mu$. Indeed, for $M_\bullet \approx 5 \times 10^8 M_\odot$, nearly half of the bias comes from such galaxies for all but the smallest $\sigma_\mu$ shown. Thus, not only

FIG. 4.—Distribution of absolute magnitude $M_V = -2.5 \log L + \text{const}$ or velocity dispersion $\log \sigma$ at selected values of $M_\bullet$ ($\log M_\bullet/M_\odot$ is given at the top right of each panel). These probability distributions are obtained by taking horizontal cuts through the joint distributions shown in Fig. 2, and are appropriate when the sample selection is on black hole mass. Two distributions are shown with dashed and solid lines, corresponding respectively to $\sigma_\mu = 0.25$, 0.50 for $M_V$ on the left or $\sigma_\mu = 0.15$, 0.30 for $\log \sigma$ on the right. The horizontal coordinate is the difference between either $M_V$ or $\log \sigma$ and the nominal values $f^{-1}(\mu)$ determined from the $M_\bullet$-L or $M_\bullet$-$\sigma$ relations, eqs. (4) and (5).
is the bias sensitive to $\sigma_\mu$, it also depends on the shape of the wings of the error distribution. If, as is likely, the wings are more extended than in our assumed log-normal distribution, the bias will be even larger.

Many of the features in these plots can be understood analytically. From the definition of $\langle s \rangle_\mu$ (eq. [11]) and equation (6), we can write

$$a + b\langle s \rangle_\mu = \mu - \frac{\int x h(x) g[b^{-1}(\mu - a - x)] \, dx}{\int h(x) g[b^{-1}(\mu - a - x)] \, dx}.$$  \hspace{1cm} (13)

If the cosmic scatter $\sigma_\mu^2 = \int dx x^2 h(x)$ is not too large, we can evaluate this by expanding $g(s)$ in a Taylor series,

$$a + b\langle s \rangle_\mu = \mu + \frac{\sigma_\mu^2}{b} \left. \frac{d \ln g(s)}{ds} \right|_{s=(\mu-a)/b} + O(\sigma_\mu^4).$$ \hspace{1cm} (14)

This result shows that (1) the bias is proportional to $\sigma_\mu^2$ if $\sigma_\mu$ is not too large; (2) the bias in $\langle s \rangle_\mu$ is positive if $g(s)$ declines with $s$ and proportional to $d \ln g(s)/ds$; this is why the bias is large and negative for the most luminous or high-dispersion galaxies; and (3) the bias is near zero if $g(s)$ is constant, which corresponds to a luminosity function $dn \propto dL/L$ when $s = \log L$. The bias in Figure 5 (left) is small for low-luminosity galaxies because the assumed luminosity function is close to this form.

A different statistical method, which more closely parallels the approach used in most observational papers, is simply to estimate the average value of the difference between $\mu$ and the prediction of the $\mu$-$s$ relation (6) for all the galaxies in the sample,

$$\Delta \log M_* \equiv \langle \mu - f(s) \rangle = \langle \mu \rangle - a - b \langle s \rangle.$$ \hspace{1cm} (15)

Let us assume that the survey flux limit at the given redshift corresponds to a luminosity $L_0$ with $\ell_0 = \log L_0$. Then,

$$\Delta \log M_* = \frac{\int_{\ell_0}^{\infty} d\ell \int g(s) ds \int (\mu - a - bs) h(\mu - a - bs) \psi(\lambda - \mu) \, d\mu}{\int_{\ell_0}^{\infty} d\ell \int g(s) ds \int h(\mu - a - bs) \psi(\lambda - \mu) \, d\mu}$$

$$= \frac{\int_{\ell_0}^{\infty} d\ell \int g(s) ds \int xh(x) \psi(\lambda - x - a - bs) \, dx}{\int_{\ell_0}^{\infty} d\ell \int g(s) ds \int h(x) \psi(\lambda - x - a - bs) \, dx}$$

$$= \frac{\int_{\ell_0}^{\infty} d\ell \int xh(x) \phi(\lambda - x) \, dx}{\int_{\ell_0}^{\infty} d\ell \int h(x) \phi(\lambda - x) \, dx},$$ \hspace{1cm} (16)

where $x \equiv \mu - a - bs$, and

$$\phi(y) = \int g(s) \psi(y - a - bs) \, ds.$$ \hspace{1cm} (17)

The luminosity function (number per unit volume) of AGNs is $dn = n\Psi(\lambda) d\lambda$, where $n$ is the number density and

$$\Psi(\lambda) = \int g(s) ds \int h(\mu - a - bs) \psi(\lambda - \mu) \, d\mu$$

$$= \int h(x) \phi(\lambda - x) \, dx,$$ \hspace{1cm} (18)

which is just the inner integral of the denominator of equation (18). To evaluate the inner integral in the numerator of equation (18), note that the functional form of $h(x)$ (eq. [7]) implies that $h'(x) = -xh(x)/\sigma_\mu^2$, so

$$\int x h(x) \phi(\lambda - x) \, dx = -\sigma_\mu^2 \int h'(x) \phi(\lambda - x) \, dx.$$ \hspace{1cm} (19)
Integrating the right side of the equation by parts, and noting that \( \lim_{y \to \infty} \phi(y) = 0 \), gives
\[
\int x h(x) \phi(\lambda - x) \, dx = \sigma_\mu^2 \int h(x) \frac{\partial}{\partial \lambda} \phi(\lambda - x) \, dx
\]
\[
= -\sigma_\mu^2 \int h(x) \frac{\partial}{\partial \lambda} \phi(\lambda - x) \, dx
\]
\[
= -\sigma_\mu^2 \frac{d}{d\lambda} \int h(x) \phi(\lambda - x) \, dx
\]
Thus,
\[
\Delta \log M_\bullet = \sigma_\mu^2 \frac{\Psi(\lambda_0)}{\int_{\lambda_0}^{\infty} \Psi(\lambda) \, d\lambda}
\]
(25)

Remarkably, the result depends only on the directly observable luminosity function \( \Psi(\lambda) \) for the type of AGN targeted in the survey and is independent of assumptions about the luminosity history \( \psi(\lambda - \mu) \) or the distribution of host galaxy properties \( g(s) \). Also, in contrast to the analogous result in equation (14), equation (25) is not just the first term in a Taylor series in \( \sigma_\mu^2 \) but valid for all values of \( \sigma_\mu^2 \), no matter how large.

To illustrate the application of this result, we adopt the quasar luminosity function from Boyle et al. (2000),
\[
\Psi(\lambda) = \frac{\Psi_*}{10^{-(1+\alpha)\lambda + \beta} + 10^{-(1+\beta)(\lambda - \lambda^*)}}
\]
(26)

with \( \alpha = -3.4 \) and \( \beta = -1.6 \), where \( \lambda^* \) is the “break” luminosity. The corresponding bias \( \Delta \log M_\bullet \) is shown as a function of the lower luminosity limit \( \lambda_0 - \lambda^* \) in Figure 6. Note that the bias in Figure 6 becomes smaller but does not vanish as the survey goes to fainter and fainter luminosity limits: asymptotically, \( \Delta \log M_\bullet \to -(1 + \beta) \ln 10 \), which equals 0.12(\( \sigma_\mu/0.3 \))^2.

3.2. Bias in a Flux-limited Survey

As a second example, we examine the bias in a flux-limited sample of AGNs. We assume that the survey galaxies are distributed uniformly in Euclidean space, that the probability distribution of AGN luminosities is given by equation (12), and that the survey contains all galaxies with flux \( f = L_{\text{AGN}}/r^2 \) exceeding some limiting flux \( f_0 \).

The bias is then given by
\[
\Delta \log M_\bullet = \int d\lambda \int g(s) nV(\lambda) \, ds
\]
\[
\times \left[ \int d\lambda \int g(s) \psi(\lambda - \mu) \, ds \right]^{-1}
\]
\[
\times \left\{ \int d\lambda \int g(s) \psi(\lambda - \mu) \, ds \right\}
\]
(27)

where \( V(\lambda) = (1/3) \Delta \Omega(L_{\text{AGN}}/f_0)^{1/2} \times 10^{33/2} \) is the volume within which an AGN of luminosity \( \lambda \) can be detected, \( \Delta \Omega \) is the solid angle covered by the survey, and \( n \) is the number density of galaxies. We have
\[
\Delta \log M_\bullet = \int d\lambda \int g(s) 10^{33/2} \, ds \int (\mu - a - bs) h(\mu - a - bs) \psi(\lambda - \mu) \, d\mu
\]
(28)
\[
= \int 10^{33/2} d\lambda \int g(s) 10^{33/2} ds \int h(\mu - a - bs) \psi(\lambda - \mu) \, d\mu
\]
(29)
\[
= \int 10^{33/2} d\lambda \int g(s) x h(x) \psi(\lambda - x - a - bs) \, dx
\]
(30)

Thus, equation (30) we use equation (24), which yields
\[
\Delta \log M_\bullet = -\sigma_\mu^2 \frac{10^{33/2} \int d\lambda [\int h(x) \psi(\lambda - x) \, dx] \, d\lambda}{\int 10^{33/2} d\lambda \int h(x) \psi(\lambda - x) \, dx}
\]
(31)
\[
= \frac{3 \ln 10}{2} - \sigma_\mu^2
\]
(32)

where the last line follows from an integration of the numerator by parts.

3.3. Selection Effects that Depend Only on Galaxy Properties

The results of the previous subsections depend on the assumption that the probability distribution of AGN luminosities is determined by the black hole mass and not the galaxy properties (eq. [12]). A foil to this is to assume that the luminosity is determined by the galaxy properties and not the black hole mass. Thus, equation (12) is replaced by
\[
dp = \chi(\lambda - ks) \, d\lambda, \text{ where } \int \chi(x) \, dx = 1
\]
(33)

where \( k \) is a constant. In this case \( \Delta \log M_\bullet = 0 \) in any flux-limited or luminosity-limited sample (Salviander et al. 2007).
4. CORRECTING FOR SELECTION BIAS MAY BE EXTREMELY DIFFICULT

Given the example bias calculations shown in the previous section, it may be tempting to conclude that similar bias corrections may be estimated and applied after the fact to existing surveys to determine the $M_*$-σ or $M_*$-L relations. We believe that in general such corrections are difficult or impossible to apply reliably to current surveys for the following reasons.

4.1. The Error Model Is Poorly Known

In the examples above, we assumed that at any galaxy property $s$, $\mu = \log M_*$ followed a normal distribution about the mean $M_*$-σ relation $\mu = f(s)$, with standard deviation characterized by a cosmic scatter $\sigma_\mu$. We saw that the bias is very sensitive to $\sigma_\mu$ (in most examples $\propto \sigma_\mu^2$), yet this parameter is poorly known for either the $M_*$-σ or $M_*$-L relations, in part because an accurate estimate of $\sigma_\mu$ requires knowing the observational errors in the $M_*$ determinations. Novak et al. (2006) have studied this problem and conclude only that $\sigma_\mu \lesssim 0.3$ for the $M_*$-σ relation and $\sigma_\mu \lesssim 0.5$ for the $M_*$-L relation. Moreover, the Novak et al. (2006) analysis assumes, as we do, that $\sigma_\mu$ is constant over the parameter range of the $M_*$ relations; the sample of galaxies with well-determined $M_*$ is simply too small to explore the possibility that it is not.

Even if an accurate estimate of $\sigma_\mu$ were available, the functional form of the distribution of $\mu - f(s)$ is unknown. The assumption of a normal distribution is an obvious first step, yet the steep fall-off of the $L$ or $\sigma$ distribution at large values means that the selection bias for large black hole masses is sensitive to the wings of this distribution, where $\mu - f(s) \approx 2\sigma_\mu$. The assumed normal distribution is likely to underestimate the selection bias, since more realistic distributions have fatter tails. Given that the local black hole sample is so small that $\sigma_\mu$ is poorly determined, the sample of galaxies with good $M_*$ determinations would have to be several orders of magnitude larger before an accurate form for the scatter function could be determined.

4.2. The $M_*$ Relationships Are Poorly Known

The $M_*$ relations are imperfectly known at $z = 0$, and in particular the range of $L$ or $\sigma$ that is used to determine them is rather limited. For example, in the sample of 31 local galaxies with measured black hole masses used by Tremaine et al. (2002) to determine the $M_*$-σ relation, the interquartile ranges in $L$ and $\sigma$ are only factors of 5.3 and 1.6, respectively. Thus, the calculation of the selection biases for $M_* \gtrsim 10^8 M_\odot$ must contend with the fact that there are only a few black holes observed in this mass range in the local sample (four in the Tremaine et al. [2002] sample). Lauer et al. (2007) showed that the $L$-σ relation must be curved (in logarithmic coordinates) in the sense that $\sigma$ appears to increase only slowly (if at all) with $L$ for the most luminous galaxies. This implies that the present log-linear $M_*$-σ and $M_*$-L relation cannot be consistent at high galaxy luminosity. The joint $M_*$-σ distributions shown in Figure 2 are thus based on extrapolated estimates at the highest $M_*$ values that are likely to change if more and better determinations of black hole masses $M_* \gtrsim 10^8 M_\odot$ become available.

4.3. What Determines AGN Luminosity?

The heart of our analysis of selection bias in §§ 3.1 and 3.2 is the assumption that the black hole determines the AGN luminosity, or more precisely, that the galaxy properties do not. This is clearly true if, for example, (1) black holes radiate either at the Eddington luminosity $L_{\text{Edd}}$ or not at all, and (2) the bolometric correction is independent of galaxy properties. This assumption is also true in more general circumstances; for example, (1) can be replaced by the weaker assumption that the probability that a black hole is radiating at some $L_{\text{AGN}}$ depends only on the ratio $L_{\text{AGN}}/L_{\text{Edd}}$ (eq. [12]). This assumption may be approximately correct; for example, Hopkins et al. (2006) found that their simulations are well fit by a model in which the probability distribution of AGN luminosities depends only on $L_{\text{AGN}}/L_{\text{peak}}$, where $L_{\text{peak}}$ is the peak luminosity of the AGN and $L_{\text{peak}} \propto L_{\text{Edd}}^{12}$. Nevertheless, if galaxy properties do affect the AGN luminosity, most likely by determining the feeding rate of matter onto the black hole at luminosities $L \ll L_{\text{Edd}}$, then the biases will be different from those presented in §§ 3.1 and 3.2. In the limiting case that the AGN luminosity is determined entirely by galaxy properties rather than black hole mass, as might be plausible for low-luminosity AGNs, there is no selection bias in a flux- or luminosity-limited survey (§ 3.3).

4.4. Did the Galaxy Make the Black Hole, or the Black Hole the Galaxy?

We have assumed that the joint probability distribution of black hole mass $\mu$ and galaxy property $s$ can be modeled by equation (2). This corresponds to a physical model in which the galaxy property is the independent variable and the black hole mass is determined by the properties of the galaxy through the cosmic-scatter function $h(y)$ and the $M_*$ relation $f(s)$. This assumption is convenient for modeling the local sample of galaxies with $M_*$ determinations, which was selected by galaxy properties, plus the ready knowledge of $q(s)$, the volume distribution of $s$. The alternative model provided by equation (3), in which the black hole mass is the independent variable and the galaxy property is determined by the black hole through the function $p(y)$, is much more difficult to fit to the local observations as we have little direct knowledge of $\Phi_q(\mu)$, the black hole mass function. However, our present understanding of black hole and galaxy formation does not allow us to say which model, if either, is correct.

4.5. Survey Selection Effects May Be Poorly Known

Most of our estimates of selection bias have been based on the assumption that surveys are complete to a given AGN luminosity or flux, but this is an oversimplification. Any survey that is based on measurements of the luminosity of the host galaxy requires that the host is bright enough and large enough to be separated from the AGN. Any survey that uses the velocity dispersion measured from stellar absorption lines requires that the absorption lines are strong enough, and that AGN emission lines in the vicinity of the absorption lines are weak enough, to allow a reliable dispersion measurement. At the opposite extreme, surveys of low-luminosity AGNs require that the emission lines are strong enough to be detected against the continuum flux from the galaxy. In such studies the selection depends in a complex way on properties such as the ratio of AGN to galaxy luminosity, and on galaxy properties other than $L$ or $\sigma$, such as the effective radius or central surface brightness.

5. BIASES IN EXISTING SURVEYS

We have intended this paper mainly as a planning guide for future surveys to explore the evolution of the $M_*$ relation, rather than as a critique of existing surveys. The following brief comments on existing surveys are mostly intended to illustrate the application and impact of the estimates of selection bias that we have made in earlier sections.

Several authors have determined the $M_*$-σ relation for nearby AGNs and compared this to the local relation for inactive galaxies.
Using seven Seyfert 1 galaxies with black hole masses measured by reverberation mapping, Gebhardt et al. (2000b) found \( \Delta \log M_\bullet = -0.21 \pm 0.13 \). For 16 AGNs with reverberation-based black hole masses Onken et al. (2004) found \( \Delta \log M_\bullet = -0.26 \pm 0.15 \) (although Onken et al. interpreted their result in terms of calibrating the reverberation-mapping method rather than an offset in the \( M_\bullet - \sigma \) relation). Nelson et al. (2004) similarly found \( \Delta \log M_\bullet = -0.26 \pm 0.46 \) from 14 AGNs, but given the error bars did not consider this offset to be significant. Barth et al. (2005) have compared the black hole masses in dwarf Seyfert 1 galaxies to an extrapolation of the \( M_\bullet - \sigma \) relation to lower dispersions; using black hole masses based on the photoionization method calibrated by the standard “isotropic” formula, they found \( \Delta \log M_\bullet = -0.04 \) (although they interpreted their results with Onken et al.’s calibration and thus found \( \Delta \log M_\bullet = +0.23 \)). Greene & Ho (2006) have measured the \( M_\bullet - \sigma \) relation in 88 nearby AGNs to determine \( M_\bullet \). Most of their masses are determined by the photoionization method, but about 20% are determined by reverberation mapping. They found \( \Delta \log M_\bullet = -0.21 \pm 0.06 \), consistent with Onken et al. (2004)—not a surprising result, since the reverberation-mapped masses that are common to the two samples have the smallest error bars and tend to dominate the fit.

Similarly, Labita et al. (2006) have determined the \( M_\bullet - L \) relation from a sample of 29 quasars with \( z < 0.6 \); they determined black hole masses using the photoionization method and luminosities using HST photometry. They found \( \Delta \log M_\bullet = -0.36 \), although they interpreted their result in terms of recalibrating the photoionization method.

A common feature of the Gebhardt et al. (2000b), Onken et al. (2004), Barth et al. (2005), and Greene & Ho (2006) samples is that their black hole masses are relatively modest: generally \( M_\bullet < 10^8 M_\odot \) and often more than an order of magnitude smaller. Selection bias occurs at all \( M_\bullet \), depending on the local slopes of the \( L \) and \( \sigma \) distributions, but these samples are safely removed from the very strong biases seen at large values of \( L, \sigma \), and \( M_\bullet \), where the distribution functions are falling rapidly. Although we have emphasized selection bias in the context of low- and high-redshift samples with similar black hole masses, one selected by galaxy properties and one by black hole mass, selection bias can also arise if the two samples are selected in the same way but are centered on different black hole mass ranges.

We further note that selection effects in these samples are difficult to model; for example, (1) as Gebhardt et al. (2000b) pointed out, the AGN variability timescale depends on luminosity, and if the timescale is too long, reverberation mapping is impractical; (2) to be able to measure the velocity dispersion requires some minimum ratio of the luminosity of the host galaxy to the luminosity of the AGN. If we make the simplest possible assumption, that the samples of nearby galaxies are complete but flux-limited, then equation (32) implies \( \Delta \log M_\bullet = 0.31 (\sigma_\mu / 0.3)^2 \).

Woo et al. (2006), building on the initial work of Treu et al. (2004), examine a sample of 14 Seyfert 1 galaxies at \( z = 0.36 \pm 0.01 \), measuring the velocity dispersions from stellar absorption lines and the black hole masses by the photoionization method. Compared to the local sample of inactive galaxies with measured black hole masses, they found \( \Delta \log M_\bullet = 0.62 \pm 0.10 \pm 0.25 \); note that this result is based on a calibration of the photoionization masses that assumes that the Onken et al. (2004) sample should have exactly the same \( M_\bullet - \sigma \) relation as inactive galaxies, which need not be the case if selection bias is accounted for. Woo et al. (2006) argued that their offset \( \Delta \log M_\bullet \) is robust because the black hole masses are measured by the same technique, with the same calibration, in the local and high-redshift samples, but the selection bias in the two samples is likely to be quite different; the Onken et al. (2004) sample is more nearly flux-limited, while the Woo et al. (2006) sample is more nearly luminosity-limited, and the selection bias in these two cases is quite different (eqns. [32] and [25]). A useful next step in assessing the bias would be to determine the luminosity function for AGNs selected by the criteria used by Woo et al. (2006) for use in equation (25).

It may also be significant that the typical \( M_\bullet \approx 4 \times 10^7 M_\odot \) in the Onken et al. (2004) sample is roughly an order of magnitude less massive than the typical \( M_\bullet \approx 3 \times 10^8 M_\odot \) in the Woo et al. (2006) sample. As is evident in Figure 5, the bias is a strong function of \( M_\bullet \), and indeed the bias in the \( M_\bullet - \sigma \) relation changes sign at \( M_\bullet \sim 10^8 M_\odot \). Thus, we would expect that selection bias might affect the results even if the Onken et al. (2004) sample were chosen in precisely the same way as the Woo et al. (2006) sample.

Figure 5 predicts a bias of \( \Delta \log M_\bullet \sim 0.2 \) at \( M_\bullet \approx 3 \times 10^8 M_\odot \) for \( \sigma_\mu = 0.30 \). This is only 40% of the offset claimed by Woo et al. (2006). Thus, it appears that selection bias cannot explain all or even most of the Woo et al. (2006) \( \Delta \log M_\bullet \); however, given the uncertainties discussed in the previous section, the possibility that the correction may be as large as the claimed offset cannot be ruled out.

Peng et al. (2006a) and Peng et al. (2006b) found \( |\Delta \log M_\bullet| \leq 0.3 \) compared to the local \( M_\bullet - L \) relation, using a total sample of 51 quasars at \( z > 1 \); the black hole masses are determined with the photoionization method. If we assume that the velocity dispersion of the host galaxy does not change between \( z \sim 2 \) and \( z = 0 \), and that the luminosity of the host fades by 1–2 mag as predicted by passive stellar evolution, then this result corresponds to \( \Delta \log M_\bullet \approx -0.04 \) as measured by the \( M_\bullet - \sigma \) relation. However, the potential selection bias is rather large for this sample, partly because the cosmic scatter in the \( M_\bullet - L \) relation may be larger than in the \( M_\bullet - \sigma \) relation. If the sample is luminosity limited, the model shown in Figure 5 (left) predicts \( \Delta \log M_\bullet = (0.3–0.7) \) for \( \sigma_\mu = 0.5 \) for the range of black hole masses in the sample, and the flux-limited model predicts \( \Delta \log M_\bullet = 0.58 \) for \( \sigma_\mu = 0.5 \) (note that our assumption of a Euclidean metric is not valid at \( z \approx 2 \), and should be generalized to a realistic cosmological model). A further complication is that their method requires measuring the bulge luminosity in the presence of the bright AGN, so imposes a selection on the ratio of AGN to galaxy luminosity. Without an accurate estimate of the selection biases for the Peng et al. (2006a) and Peng et al. (2006b) samples, their measurements of \( \Delta \log M_\bullet \) do not provide clear evidence for rapid evolution in the ratio of black hole mass to stellar mass in the galaxy.

Salviander et al. (2007) found that galaxies of a given dispersion at \( z \approx 1 \) have black hole masses that are larger by \( \Delta \log M_\bullet \sim 0.45 \) than at \( z \approx 0 \). In contrast to most other studies, they carefully consider the effects of selection bias using Monte Carlo models and conclude that for \( \sigma_\mu = 0.3 \) selection bias increases \( \log M_\bullet \) by \( \sim 0.1 \) at low redshift and 0.2 at \( z \approx 1 \), contributing a net offset of 0.1. They also argue that scatter in the relationship between the width of the [O iii] emission line and the true stellar dispersion contributes an additional offset of \( \sim 0.15 \), so after correcting for these biases \( \Delta \log M_\bullet \sim 0.2 \). Their estimates of selection bias are similar to, but somewhat smaller than, the bias that we estimate for their sample from equations (25) and (32), assuming that the low-redshift sample is flux limited and the high-redshift sample is luminosity limited; we believe our estimates are more robust because they do not depend on an assumed
form for the distribution \(\psi(\lambda - \mu)\) of AGN luminosities at a given black hole mass.

Finally, we note that theoretical models of the evolution of the \(M_\bullet - \sigma\) relation (Robertson et al. 2006) predicted \(\Delta \log M_\bullet = -0.18\) at \(z = 2\). Strong selection bias could reduce or eliminate the difference between this prediction and the positive values of \(\Delta \log M_\bullet\) found at high redshift in most of the above studies.

6. Recapitulation

We have described a selection bias that affects investigations into the cosmological evolution of the \(M_\bullet - \sigma\) or \(M_\bullet - L\) relations. We summarize the main points as follows:

1. Cosmic scatter in the \(M_\bullet - \sigma\) or \(M_\bullet - L\) relations means that a galaxy of given \(L\) or \(\sigma\) will host black holes with a range of \(M_\bullet\) values.

2. The steep decline in the luminosity function for the most luminous galaxies means that the rare occurrence of high-mass black holes in numerous "modest" galaxies overwhelms the frequent occurrence of black holes of similar mass in the rare galaxies with very high luminosity. A similar conclusion applies to dispersion instead of luminosity.

3. Selection of a sample of black holes hosted by inactive galaxies at low redshift explores the distribution of black hole masses \(M_\bullet\) that a galaxy will host, given its luminosity \(L\) or dispersion \(\sigma\). Selection of a sample by AGN luminosity at high redshift explores the distribution of \(L\) or \(\sigma\) at a given \(M_\bullet\) (if, as is likely, the distribution of AGN luminosity is determined mainly by the black hole mass rather than the galaxy properties). Confusing the two distributions may create a false signature of evolution.

4. The selection bias is substantial for the current estimates of the cosmic scatter in the \(M_\bullet - L\) or \(M_\bullet - \sigma\) relations, particularly for galaxies containing the most massive black holes. Neither the rms cosmic scatter nor the shape of the scatter function is known accurately, and hence the bias cannot be reliably corrected for, given our present and even foreseeable state of knowledge.

5. Our rough estimates of selection bias as applied to a number of evolutionary studies show that the potential bias may be larger than many of the \(\Delta \log M_\bullet\) or \(\Delta s\) values observed. With the notable exceptions of Salviander et al. (2007) and Treu et al. (2007) it appears that most studies of the evolution of the \(M_\bullet - \sigma\) or \(M_\bullet - L\) relations have not considered the effects of selection bias. In most cases, accounting properly for selection bias is likely to reduce the evolution in the \(M_\bullet\) relations that has been observed in recent surveys.

6. The only way to avoid selection bias is to choose high- and low-redshift galaxy samples using precisely defined, objective criteria that are precisely the same for the two samples. To our knowledge, this has not yet been done.

This work was initiated during a Lorentz Center workshop attended by the authors. We thank the staff of the Lorentz Center and Tim de Zeeuw and Richard McDermid for hosting us. We thank Karl Gebhardt, Richard Green, Jenny Greene, Tommaso Treu, and Nadia Zakamska for useful conversations. We thank the referee for a number of suggestions that improved characterization of the evolutionary investigations cited in the text.

REFERENCES

Barth, A. J., Greene, J. E., & Ho, L. C. 2005, ApJ, 619, L151
Barth, A. J., Greene, J. E., Ho, L. C., 2006, AJ, 131, 1018
Blanton, M. R., et al. 2003, ApJ, 592, 819
Boyle, B. J., Shanks, T., Croom, S. M., Smith, R. J., Miller, L., Loaring, N., & Heymans, C. 2000, MNRAS, 317, 1014
Dressler, A. 1989, in IAU Symp. 134, Active Galactic Nuclei, ed. D. Osterbrock & J. S. Miller (Dordrecht: Kluwer), 217
Ferrarese, L., & Merritt, D. 2000, ApJ, 539, L9
Fine, S., et al. 2006, MNRAS, 373, 613
Gebhardt, K., et al. 2000a, ApJ, 539, L13
Gebhardt, K., et al. 2000b, ApJ, 543, L5
Greene, J. E., & Ho, L. C. 2006, ApJ, 641, L21
H"aring, N., & Rix, H. 2004, ApJ, 604, L89
Hopkins, P. F., Hernquist, L., Cox, T. J., Di Matteo, T., Robertson, B., & Springel, V. 2006, ApJS, 163, 1
Kormendy, J. 1993, in The Nearest Active Galaxies, ed. J. Beckman, L. Colina, & H. Netzer (Madrid: CSIC), 197
Kormendy, J., & Richstone, D. 1995, ARA&A, 33, 581
Labita, M., Treves, A., Falomo, R., & Uslenghi, M. 2006, MNRAS, 373, 551
Lauer, T. R., et al. 2007, ApJ, 662, 808
Magorrian, J., et al. 1998, AJ, 115, 2285
Malumquist, K. G. 1924, Medd. Lund Astron. Obs. II, 32, 64
Nelson, C. H., Green, R. F., Bower, G., Gebhardt, K., & Weistrop, D. 2004, ApJ, 615, 652
Novak, G. S., Faber, S. M., & Dekel, A. 2006, ApJ, 637, 96
Onken, C. A., Ferrarese, L., Merritt, D., Peterson, B. M., Pogge, R. W., Vestergaard, M., & Wandell, A. 2004, ApJ, 615, 645
Peng, C. Y., Impey, C. D., Ho, L. C., Barton, E. J., & Rix, H.-W. 2006, ApJ, 640, 114
Peng, C. Y., Impey, C. D., Rix, H.-W., Koachane, C. S., Keeton, C. R., Falco, E. E., Leh"ar, J., & McLeod, B. A. 2006, ApJ, 649, 616
Postman, M., & Lauer, T. R. 1995, ApJ, 440, 28
Robertson, B., Hernquist, L., Cox, T. J., Di Matteo, T., Hopkins, P. F., Martini, P., & Springel, V. 2006, ApJ, 641, 90
Salviander, S., Shields, G. A., Gebhardt, K., & Bonning, E. W. 2007, ApJ, 662, 131
Schechter, P. 1976, ApJ, 203, 297
Sheth, R. K., et al. 2003, ApJ, 594, 225
Shields, G. A., Gebhardt, K., Salviander, S., Wills, B. J., Xie, B., Brotherton, M. S., Yuan, J., & Dietrich, M. 2003, ApJ, 583, 124
Tremaine, S., et al. 2002, ApJ, 574, 740
Treu, T., Malkan, M. A., & Blandford, R. D. 2004, ApJ, 615, L97
Treu, T., Woo, J.-H., Malkan, M. A., & Blandford, R. D. 2007, ApJ, 667, 117
Tundo, E., Bernardi, M., Hyde, J. B., Sheth, R. K., & Pizzella, A. 2007, ApJ, 663, 53
Vestergaard, M. 2006, NewA Rev., 50, 817
Yu, Q., & Lu, Y. 2004, ApJ, 602, 603
Yu, Q., & Tremaine, S. 2002, MNRAS, 335, 965
Willott, C. J., Percival, W. J., McLure, R. J., Crampton, D., Hutchings, J. B., Jarvis, M. J., Sawicki, M., & Simard, L. 2005, ApJ, 626, 657
Woo, J.-H., Treu, T., Malkan, M. A., & Blandford, R. D. 2006, ApJ, 645, 900