The Astrophysical Journal, 688:48–58, 2008 November 20
© 2008. The American Astronomical Society. All rights reserved. Printed in U.S.A.

RECENT STRUCTURAL EVOLUTION OF EARLY-TYPE GALAXIES: SIZE GROWTH FROM $z = 1$ TO $z = 0$[1]

Arjen van der Wel
Department of Physics and Astronomy, Johns Hopkins University, 3400 North Charles Street, Baltimore, MD 21218; wel@pha.jhu.edu

Bradford P. Holden
University of California Observatories/Lick Observatory, University of California, Santa Cruz, CA 95064

Andrew W. Zirm
Department of Physics and Astronomy, Johns Hopkins University, 3400 North Charles Street, Baltimore, MD 21218

Marijn Franx
Leiden Observatory, Leiden University, P.O. Box 9513, NL-2300 AA Leiden, Netherlands

Alessandro Rettura
Department of Physics and Astronomy, Johns Hopkins University, 3400 North Charles Street, Baltimore, MD 21218

Garth D. Illingworth
University of California Observatories/Lick Observatory, University of California, Santa Cruz, CA 95064

and

Holland C. Ford
Department of Physics and Astronomy, Johns Hopkins University, 3400 North Charles Street, Baltimore, MD 21218
Received 2008 June 14; accepted 2008 July 31

ABSTRACT

Strong size and internal density evolution of early-type galaxies between $z \sim 2$ and the present has been reported by several authors. Here we analyze samples of nearby and distant ($z \sim 1$) galaxies with dynamically measured masses in order to confirm the previous, model-dependent results and constrain the uncertainties that may play a role. Velocity dispersion (σ) measurements are taken from the literature for 50 morphologically selected $0.8 < z < 1.2$ field and cluster early-type galaxies with typical masses $M_{\rm dyn} = 2 \times 10^{11}\ M_{\odot}$. Sizes ($R_{\rm eff}$) are determined with Advanced Camera for Surveys imaging. We compare the distant sample with a large sample of nearby ($0.04 < z < 0.08$) early-type galaxies extracted from the Sloan Digital Sky Survey for which we determine sizes, masses, and densities in a consistent manner, using simulations to quantify systematic differences between the size measurements of nearby and distant galaxies. We find a highly significant difference between the σ-$R_{\rm eff}$ distributions of the nearby and distant samples, regardless of sample selection effects. The implied evolution in $R_{\rm eff}$ at fixed mass between $z = 1$ and the present is a factor of 1.97 ± 0.15. This is in qualitative agreement with semianalytic models; however, the observed evolution is much faster than the predicted evolution. Our results reinforce and are quantitatively consistent with previous, photometric studies that found size evolution of up to a factor of 5 since $z \sim 2$. A combination of structural evolution of individual galaxies through the accretion of companions and the continuous formation of early-type galaxies through increasingly gas-poor mergers is one plausible explanation of the observations.

Subject headings: galaxies: clusters: general — galaxies: elliptical and lenticular, cD — galaxies: evolution — galaxies: formation — galaxies: fundamental parameters — galaxies: general — galaxies: photometry

Online material: color figures

1. INTRODUCTION

Hierarchical galaxy formation models embedded in a ΛCDM cosmology predict strong size evolution for massive galaxies. A higher gas fraction in high-redshift galaxies leads to more dissipation and hence compact galaxies (e.g., Robertson et al. 2006; Khochfar & Silk 2006b), and subsequent evolution such as dry merging or accretion of smaller systems can increase the size of a galaxy (e.g., Loeb & Peebles 2003; Naab et al. 2007). The models predict the strongest sample-averaged size evolution for the most massive galaxies (Khochfar & Silk 2006a) because of large differences in the gas fraction at different redshifts and because the assembly of massive galaxies continues until very late epochs in a hierarchical framework (e.g., De Lucia et al. 2006).

Evidence for significant size evolution between $z \sim 2$ and the present has been building up quickly over the past few years (e.g., Trujillo et al. 2004, 2006b; Franx et al. 2008). In particular, galaxies with low star formation rates and high stellar masses ($\gtrsim 10^{11}\ M_{\odot}$) appear to be extremely compact from $z \sim 1.5$ (Daddi et al. 2005; Trujillo et al. 2006a, 2007; Longhetti et al. 2007; Cimatti et al. 2008; Rettura et al. 2008) to $z \sim 2.5$ (Zirm et al. 2007; Toft et al. 2007; van Dokkum et al. 2008; Buitrago et al. 2008). Given the similarity between many of their observed

[1] Based on observations with the *Hubble Space Telescope*, obtained at the Space Telescope Science Institute, which is operated by AURA, Inc., under NASA contract NAS5-26555, and observations made with the *Spitzer Space Telescope*, which is operated by the Jet Propulsion Laboratory, California Institute of Technology, under NASA contract 1407. Based on observations collected at the European Southern Observatory, Chile (169.A-0458). Some of the data presented herein were obtained at the W. M. Keck Observatory, which is operated as a scientific partnership among the California Institute of Technology, the University of California and the National Aeronautics and Space Administration. The Observatory was made possible by the generous financial support of the W.M. Keck Foundation.

properties, there is likely to be an evolutionary connection between these distant compact galaxies and the present-day early-type galaxies despite the measured large difference of 2 orders of magnitude in surface mass density (e.g., Zirm et al. 2007).

The measurements of sizes and densities of high-redshift galaxies are hampered by many systematic uncertainties (e.g., morphological *K*-corrections, surface brightness dimming, errors in photometric redshifts, and mass measurements). Most of these errors, however, are unlikely to fully account for the observed strong size evolution. The uncertainty in the mass estimates may be the exception. For the work in the literature, these mass estimates are always based on the photometric properties of the galaxies. For a reasonable set of assumptions the photometric stellar mass estimates are not uncertain by more than a factor of 2 or 3 and would not change the inferred evolution significantly. However, since we infer that $z \sim 2$ galaxies must have physical central densities that are 3 orders of magnitude higher than those of local galaxies (van Dokkum et al. 2008), further verification of those apparently reasonable assumptions is warranted. For example, a stellar initial mass function (IMF) that is radically different (e.g., Larson 2005; Fardal et al. 2007; van Dokkum 2008; Davé 2008) from a Salpeter-like IMF (Salpeter 1955; Scalo 1986; Kroupa 2001; Chabrier 2003; Hoversten & Glazebrook 2008) could reduce the stellar mass estimates by an order of magnitude, producing perfectly normal galaxies by today's standards.

The spectacular nature of these compact galaxies at $z \sim 2$ could be confirmed by direct, kinematical mass measurements. However, the quiescent nature of these objects and their consequent lack of emission lines (Kriek et al. 2006) require absorption-line measurements of their stellar velocity dispersions, which should be as high as 400–500 km s^{-1} (Toft et al. 2007; van Dokkum et al. 2008). Unfortunately, with the currently available instrumentation this is not feasible. These $z \sim 2$ galaxies are prohibitively faint at observed optical wavelengths (see, e.g., Cimatti et al. 2008), and near-infrared spectroscopy is still maturing as a technique. Continuum detections in the observed NIR have only recently become possible (Kriek et al. 2006) for the brightest sources, and no detection of absorption lines has been made.

At lower redshifts ($z \sim 1$) absorption-line spectroscopy has for years been a powerful tool to study the evolution of distant early-type galaxies (van Dokkum et al. 1998; van Dokkum & Stanford 2003; van Dokkum & Ellis 2003; Wuyts et al. 2004; van der Wel et al. 2004; Treu et al. 2005a, 2005b; Holden et al. 2005; van der Wel et al. 2005; di Serego Alighieri et al. 2005; Jørgensen et al. 2005). Size evolution is a gradual process (see, e.g., Trujillo et al. 2006b); therefore, intermediate changes in sizes and densities should be observable at these redshifts.

In this paper we compile a sample of galaxies at redshifts $0.8 < z < 1.2$ with measured absorption-line velocity dispersions from the literature and that are visually classified as early-type galaxies with the aid of *Hubble Space Telescope* (*HST*) imaging from the Advanced Camera for Surveys (ACS; Ford et al. 1998). We measure the galaxies' sizes from these ACS data. We then compare this distant sample with nearby early-type galaxies extracted from the Sloan Digital Sky Survey (SDSS; York et al. 2000). This comparison, with careful control of systematic uncertainties, allows us to verify that distant early-type galaxies are indeed significantly more compact than their local counterparts.

The advantage of this approach is that the size and density measurements are independent of the photometric properties of the galaxies apart from the surface brightness profile. The absence of luminosity and other photometric properties from our analysis assures us that our study does not suffer from the strong possible biases in previous photometric work. Moreover, deep, high-resolution ACS imaging allows us to determine sizes of $z \sim 1$ galaxies to a precision comparable to that for nearby galaxies and in a consistent manner. Most previous studies verify for biases in the size determinations within their distant samples (e.g., Trujillo et al. 2004, 2006b; Cimatti et al. 2008) but do not extend this analysis to verify the consistency with size measurements of nearby galaxies.

In § 2 we describe the sample of nearby early-type galaxies and derive the dynamical mass-size relation. In § 3 we construct the sample of $z \sim 1$ early-type galaxies, determining their masses and sizes in a manner that is consistent with the nearby sample. In § 4 we quantify systematic effects in our size measurements through simulations. In § 5 we derive the evolution in the dynamical mass-size relation. In § 6 we compare our results with previous measurements based on photometric mass estimates and semianalytic model predictions, and we discuss size evolution in the broader context of the evolving early-type galaxy population. Finally, in § 7 we summarize our results and conclusions. We adopt the following cosmological parameters: $(\Omega_M, \Omega_\Lambda, h) = (0.3, 0.7, 0.7)$.

2. NEARBY EARLY-TYPE GALAXIES

2.1. *Velocity Dispersions and Sizes*

We have extracted a large sample of early-type galaxies at redshifts $0.04 < z < 0.08$ from the SDSS database (DR6; Adelman-McCarthy et al. 2008) based on the criteria as outlined by Graves et al. (2007).[2] Briefly, galaxies on the red sequence and either without emission lines or with high [O II] to Hα ratios are included in the sample. These criteria effectively exclude star-forming galaxies but include genuine early-type galaxies with nuclear activity (see Yan et al. 2006).

The dispersion as measured within the spectroscopic aperture ($\sigma_{\rm ap}$) is corrected to match the average dispersion within the effective radius $R_{\rm eff}$ (measured as described below) following Jørgensen et al. (1995):

$$\sigma_{\rm eff} = \sigma_{\rm ap} \left(\frac{R_{\rm eff}}{R_{\rm ap}(z)} \right)^{-0.04}, \quad (1)$$

where $R_{\rm ap}(z)$ is the radius of the SDSS spectroscopic fiber (1.5″) in kpc at the distance of the galaxy. We use the correction from Jørgensen et al. (1995) for consistency with previously published results. We note that Cappellari et al. (2006) used better data to improve the aperture correction, but the resulting difference in $\sigma_{\rm eff}$ is only a few percent.

We use GALFIT (Peng et al. 2002) to determine effective radii from the SDSS *g*-band imaging assuming an $R^{1/4}$ profile, leaving the effective radius, the integrated magnitude, the position angle, the axial ratio, and the position of the center as free parameters. The point-spread function (PSF), which is used to deconvolve the image, is constructed for each galaxy separately by co-adding the stars in the frames after drizzling the cutouts to a common center. A more general $R^{1/n}$ profile (Sérsic 1968) may provide a more realistic description of the surface brightness distribution of individual early-type galaxies, especially in the presence of a significant disk. However, $n = 4$ provides a good description of the average profile of early-type galaxies both nearby (de Vaucouleurs 1948) and at $z \sim 1$ (see § 3.2). Moreover, introducing *n* as an additional free parameter results in unnecessarily large, redshift-dependent systematic uncertainties in the size measurements (see § 4).

[2] The IDs, positions, redshifts, and dispersions for this sample were kindly provided by G. Graves.

The size parameter that we use in this paper is the circularized effective radius $\sqrt{ab} \equiv a\sqrt{q}$, where a is the effective radius along the major axis (the output parameter of GALFIT), b is the effective radius along the minor axis, and q the axis ratio (as calculated by GALFIT); $\sqrt{ab}$ is a good approximation for optically thin luminosity distributions such as the generally dust-poor early-type galaxies in our samples. The systematic and random errors of our size determinations are inferred from extensive simulations described in § 4.

The SDSS spectroscopic catalog suffers from several biases that may mitigate size evolution measurements. First, compact sources are not targeted for spectroscopy as they may be mistaken for stars or because their central surface brightnesses, i.e., their fiber magnitudes, are too bright. Second, almost all galaxies that in the literature (see the HyperLEDA database compiled by Paturel et al. 2003) have been claimed to have high, >300 km s^{-1} velocity dispersions have dispersions of <300 km s^{-1} in the SDSS (see Bernardi 2007, Appendix A). The source of this discrepancy is unknown. While Bernardi convincingly argues that the SDSS dispersions are more reliable, there are a number of galaxies with large, mutually consistent dispersion measurements from multiple, independent observers, and for which the SDSS dispersion is significantly smaller. These potential caveats may cause our size evolution measurements to be biased. We refer to these issues when we present our results in § 5.

2.2. *The Mass-Radius Relation and the Mass-Density Relation*

From $\sigma_{\rm eff}$ and $R_{\rm eff}$ we derive the total dynamical mass and the corresponding average surface mass density within $R_{\rm eff}$:

$$M_{\rm dyn} = \frac{\beta R_{\rm eff}\sigma_{\rm eff}^2}{G}, \quad (2)$$

$$\Sigma_{\rm eff} = \frac{\beta\sigma_{\rm eff}^2}{2\pi G R_{\rm eff}}, \quad (3)$$

with $\beta = 5$, which has been shown to hold for local galaxies (Cappellari et al. 2006). Following Shen et al. (2003) we adopt the following characterization of the $M_{\rm dyn}$-$R_{\rm eff}$ relation:

$$R_{\rm eff} = R_c\left(\frac{M_{\rm dyn}}{M_c}\right)^b. \quad (4)$$

With a least-squares linear fit to all galaxies with mass $M_{\rm dyn} > 3 \times 10^{10}\ M_\odot$ we find that the slope is $b = 0.56$ and the zero point normalized to a characteristic mass $M_c = 2 \times 10^{11}\ M_\odot$ is $R_c = 4.80$ kpc. We find statistically the same relation if we perform a linear fit to the values of the median $R_{\rm eff}$ in 0.1 dex wide mass bins in the range $10.5 < \log(M_{\rm dyn}) < 12.1$. The scatter around the best-fit relation is 0.14 dex.

Using stellar masses, M_*, Shen et al. (2003) find the same slope $b = 0.56$ for the M_*-$R_{\rm eff}$ relation. Their zero point, however, is larger ($R_c = 6.14$ kpc). This is likely due to the difference between $M_{\rm dyn}$ and M_* as Cappellari et al. (2006) show for a Kroupa (2001) IMF (which is also used by Shen et al. 2003) that $M_{\rm dyn} \sim 1.4 M_*$. This translates into a difference in R_c of ∼20%, close to the observed difference. Furthermore, Shen et al. (2003) analyze SDSS r-band imaging whereas we use g-band imaging.

3. DISTANT EARLY-TYPE GALAXIES

3.1. *Velocity Dispersions and Sizes*

Several authors have published velocity dispersion measurements of early-type galaxies at $z \sim 1$. We compile the data from three different data sets for which the selection criteria are well understood so that systematic effects introduced through selection effects can be properly modeled. Our compiled sample contains galaxies in the redshift range $0.8 < z < 1.2$ in the Chandra Deep Field–South (CDF-S; van der Wel et al. 2004, 2005) and the Hubble Deep Field–North (HDF-N; Treu et al. 2005a, 2005b). In addition, we include galaxies in the massive, X-ray–selected cluster MS 1054−0321 at $z = 0.831$ (Wuyts et al. 2004). The seven cluster galaxies at $z > 1$ for which dispersions have been measured (van Dokkum & Stanford 2003; Holden et al. 2005) are not included because of the paucity of this sample, which prevents us to accurately model selection effects. The final sample only contains galaxies with S/N > 10 Å^{-1} since dispersions derived from spectra with lesser quality can suffer from large (>10%) systematic uncertainties. The same aperture corrections are included as for the nearby galaxies (eq. [1]), with in this case $R_{\rm ap}(z)$ the radius of a circle with area $1'' \times 1.25''$ (the width of the slits and the typical length of the extracted region) in kiloparsecs at the redshift of the galaxy. The data are given in Table 1.

For all galaxies ACS imaging is available. In order to produce an internally consistent data set, we remeasure the sizes of all galaxies. GOODS[3] provides deep, publicly available ACS imaging of the CDF-S and the HDF-N (Giavalisco et al. 2004) in four filters. We use the F850LP (z_{850}-band) images in order to match the rest-frame wavelength at which the sizes of the nearby comparison sample are measured (the SDSS g-band; see § 2). For the MS 1054−0321 cluster ACS imaging has been taken as part of the guaranteed time observation program (Blakeslee et al. 2006). We use the F775W (i_{775}-band) imaging as the available z-band imaging is of lesser quality. At this redshift rest frame g falls in between i_{775} and z_{850} such that the morphological K-correction is not a problem; the $z \sim 0.8$ galaxies in the sample of Treu et al. (2005b) are only 3% ± 4% smaller in the i_{775} band than in the z_{850} band.

With GALFIT we determine effective radii in the same manner as those for the nearby galaxy sample (see § 2.1). The PSF is constructed with Tiny Tim (Krist 1995), even though using stars results in negligible differences (see, e.g., van der Wel et al. 2005; Treu et al. 2005b). Errors are discussed in § 4, and the data are given in Table 1.

We use equations (2) and (3) to compute masses and surface densities. Again, we adopt $\beta = 5$, which has been shown to hold for distant nonrotating elliptical galaxies (van der Marel & van Dokkum 2007; van der Wel & van der Marel 2008). For rotating early-type galaxies the situation appears to be more complex (van der Wel & van der Marel 2008) in the sense that β is possibly ∼20% larger than 5. We comment on the impact of this possible complication on our size evolution measurement in § 5.2. A low-mass cutoff of $3 \times 10^{10}\ M_\odot$ is applied since below this limit no useful samples are available due to severe incompleteness of the surveys (see, e.g., van der Wel et al. 2005).

3.2. *The Average Surface Brightness Profile*

In determining the sizes of the nearby and distant galaxies in the previous sections we assumed that an $R^{1/4}$ profile provides an accurate description of early-type galaxies. We know this to be true for nearby galaxies, but not for $z \sim 1$ early-type galaxies.

If a more general $R^{1/n}$ profile is adopted, measured values tend to cluster around $n = 4$ (see, e.g., Blakeslee et al. 2006; Rettura et al. 2006). However, there is a possibility that the true values of n are different; at large radii the "wings" of the profile become quickly overwhelmed by background noise, even in the deepest

[3] See http://www.stsci.edu/science/goods/.

TABLE 1
Velocity Dispersions and Sizes of the Distant Sample

ID	z	$\sigma_{\rm eff}$	$R_{\rm eff}$
MS 1054−1649	0.831	243 ± 28	4.91
MS 1054−2409	0.831	287 ± 33	3.30
MS 1054−3058	0.831	303 ± 33	10.20
MS 1054−3768	0.831	222 ± 24	3.28
MS 1054−3910	0.831	295 ± 42	1.80
MS 1054−4345	0.831	336 ± 34	4.35
MS 1054−4520	0.831	322 ± 30	15.20
MS 1054−4705	0.831	253 ± 36	8.84
MS 1054−4926	0.831	310 ± 38	2.04
MS 1054−5280	0.831	259 ± 31	3.68
MS 1054−5298	0.831	284 ± 39	3.54
MS 1054−5347	0.831	254 ± 24	2.94
MS 1054−5450	0.831	234 ± 26	8.16
MS 1054−5529	0.831	182 ± 23	3.24
MS 1054−5577	0.831	305 ± 40	2.67
MS 1054−5666	0.831	286 ± 23	4.99
MS 1054−5756	0.831	232 ± 27	3.98
MS 1054−6036	0.831	254 ± 22	2.93
MS 1054−6301	0.831	249 ± 24	3.55
MS 1054−6688	0.831	274 ± 37	2.93
HDFN−206	0.936	199 ± 18	1.11
HDFN−237	0.851	280 ± 21	1.80
HDFN−256	0.974	306 ± 14	3.06
HDFN−635	0.820	201 ± 17	2.29
HDFN−681	0.842	341 ± 30	1.43
HDFN−761	1.013	374 ± 39	3.77
HDFN−811	0.848	216 ± 14	1.32
HDFN−933	0.847	305 ± 37	1.86
HDFN−951	0.854	235 ± 17	2.15
HDFN−1236	0.850	217 ± 12	1.98
HDFN−1286	0.846	247 ± 17	3.66
HDFN−1287	0.846	342 ± 23	3.94
HDFN−1328	0.845	250 ± 34	1.82
HDFN−1543	0.849	280 ± 13	1.52
HDFN−1559	0.943	178 ± 12	1.66
HDFN−1633	0.841	330 ± 13	1.96
HDFN−1706	0.913	215 ± 12	2.23
HDFN−1709	0.842	218 ± 11	1.16
CDFS−1	1.089	231 ± 16	2.83
CDFS−2	0.964	200 ± 10	2.30
CDFS−3	1.044	300 ± 32	1.00
CDFS−4	0.964	336 ± 19	6.84
CDFS−7	1.135	232 ± 20	5.77
CDFS−12	1.123	262 ± 21	0.94
CDFS−13	0.980	247 ± 11	2.20
CDFS−14	0.984	197 ± 23	2.80
CDFS−18	1.096	324 ± 36	3.97
CDFS−20	1.022	199 ± 16	2.24
CDFS−25	0.967	258 ± 19	0.86
CDFS−29	1.128	221 ± 18	1.59

Notes.—The IDs and velocity dispersions are taken from Wuyts et al. (2004), Treu et al. (2005b), and van der Wel et al. (2005) for MS 1054−0321, HDFN, and CDFS, respectively. The dispersions are all aperture-corrected according to eq. (1). The sizes are determined as described in § 3.1, and a standard error of 14% is adopted for all galaxies, which is based on the simulations described in § 4.

HST imaging. Because n and the measured $R_{\rm eff}$ are correlated, assuming $n = 4$ for all redshifts introduces systematic errors in case n evolves with redshift.

In order to examine the profiles of the $z \sim 1$ galaxies at large radii, we median-stack the z_{850}-band images of all elliptical galaxies (S0s are excluded) without bright neighbors in our CDF-S and HDF-N samples (see Fig. 1). The images of the individual galaxies are drizzled onto a common central position. Due to imperfections in this procedure, the stacked PSF may not be an accurate description of the PSF of the stacked image. However, this does not play a role since the deviations we are interested in have scales that are an order of magnitude larger than the PSF.

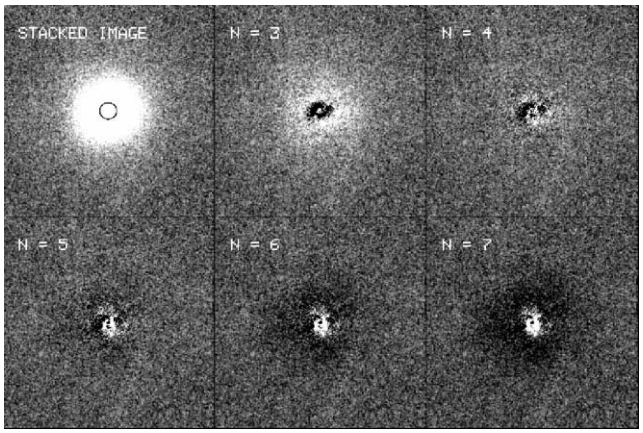

Fig. 1.—Stacked z_{850}-band image of 29 $z \sim 1$ elliptical galaxies in the CDF-S and the HDF-N and the residuals after subtracting $R^{1/n}$ model profiles, with $n = 3$, 4, 5, 6, 7. Each panel is 7.68″ on a side, which corresponds to 62 kpc at $z = 1$. The circle indicates the effective radius as measured with the $R^{1/4}$ model profile (0.31″, or 2.5 kpc at $z = 1$). The model with $n = 4$ provides the best fit. The model with $n = 3$ produces positive residuals at large radii; models with high n produce negative residuals. This justifies our choice to adopt the $R^{1/4}$ law to model the surface brightness profiles of both local and distant early-type galaxies.

With GALFIT we subtract $R^{1/n}$ profiles with integer values $n = 3$–7 (see Fig. 1). The negative residuals outside $R_{\rm eff}$ for models with large n and the positive residuals for models with small n indicate that these limiting cases provide poor fits of the outer regions of elliptical galaxies at $z \sim 1$. The $R^{1/4}$ and $R^{1/5}$ profiles provide the best description of their average surface brightness distributions. This visual impression is confirmed by the χ^2-values of the respective fits: $\chi^2 = 0.5$ for both $n = 4$ and for $n = 5$, whereas $\chi^2 > 0.7$ for other values of n. Interestingly, n does not evolve significantly with redshift, and we conclude that it is safe to assume that choosing $n = 4$ for both nearby and distant early-type galaxies does not introduce significant systematic errors.

4. SIMULATIONS OF SIZE MEASUREMENTS

To test the robustness of our size determinations of local and distant early-type galaxies in §§ 2 and 3 we simulate size measurements by using SDSS g-band imaging of 45 early-type galaxies in the Virgo Cluster (Mei et al. 2007). The pixels of the mosaics of the Virgo Cluster galaxies are rebinned to account for the different pixel scales of the various instruments, and different cosmological distances of the galaxies at higher redshifts. The redshift range is $z = 0.04$–0.08 for the nearby sample and 0.8–1.2 for the distant sample. The physical sizes of the simulated galaxies are thus conserved. For each redshift ($z = 0.04, 0.06, 0.08, 0.8, 1.2$) we run ~200 simulations with different values for the flux density of the simulated galaxies, which are chosen such that the simulated galaxies have the same range in apparent magnitude as the observed galaxies in our samples. After convolution with the appropriate PSF and the addition of Poisson noise, the seed galaxies are inserted into empty parts of real images. Their sizes are measured with GALFIT in the same manner, by fitting a $R^{1/4}$ law, as for the real galaxies. Because we are mainly interested in the systematic differences in the size determinations within and between our nearby and distant samples, we assume the size

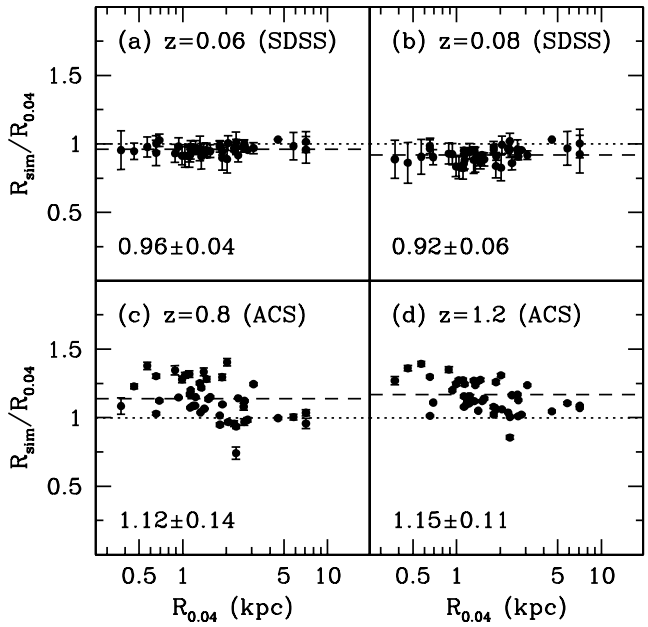

FIG. 2.—Random and systematic errors in the size determinations of early-type galaxies with SDSS imaging at (*a*) $z = 0.06$ and (*b*) $z = 0.08$ and with *HST* imaging at (*c*) $z = 0.8$ and (*d*) $z = 1.2$, all with respect to the size measurements at $= 0.04$, which are used as the benchmark in our analysis. These are the results of simulations with 45 early-type galaxies in the Virgo Cluster. The systematic offsets are indicated by the dashed lines. The scatter in the offsets, considered to be the random error in the size determinations, are also listed.

determinations based on the $z = 0.04$ simulated SDSS images of the Virgo Cluster galaxies as the baseline against which we compare the other simulated size measurements.

The results of the simulations are shown in Figure 2. Within the nearby sample we find systematic, redshift-dependent differences, of up to ~10%. Random errors, derived from the scatter in the sizes inferred from the simulated images, are typically less than 5%. Systematic differences between the nearby and distant samples can be as large as 20% for small galaxies, where at high redshift the sizes are overestimated. Random errors are typically 10%–15%. We find no systematic trends with magnitude. The reason for this is that all galaxies are relatively bright compared to the depth of the data sets, such that the limiting factor in the size measurements is spatial resolution.

Adopting a $R^{1/n}$ law with n as a free parameter may improve the quality of the fits to individual galaxies. However, our simulations reveal that the random errors increase to ~20%–25% without much change in the systematic errors. Together with the analysis of stacked images (§ 3.2), this test justifies our choice to use the $R^{1/4}$ law to describe the surface brightness profiles of all galaxies in both the nearby and the distant samples.

The sizes we derive in the §§ 2 and 3, and the derived quantities $M_{\rm dyn}$ (eq. [2]) and $\Sigma_{\rm eff}$ (eq. [3]), are corrected to account for systematic measurement errors. Those corrections depend on redshift and are interpolated between the values listed in Figure 2. For simplicity the dependence on size is not taken into account, such that the remaining systematic uncertainty is ~5%.

5. STRUCTURAL EVOLUTION OF EARLY-TYPE GALAXIES

5.1. *Evolution of the gma-Radius Distribution*

In Figure 3 we compare the $\sigma_{\rm eff}$-$R_{\rm eff}$ distributions of the nearby and distant early-type galaxy samples. This unusual projection of the fundamental plane (FP; Dressler et al. 1987; Djorgovski & Davis 1987) has a very large scatter. However, the advantage is that changes with redshift are independent of luminosity evolution. Despite the large scatter, it is clear that the distant galaxies are offset from the nearby galaxies. Galaxies with the properties of

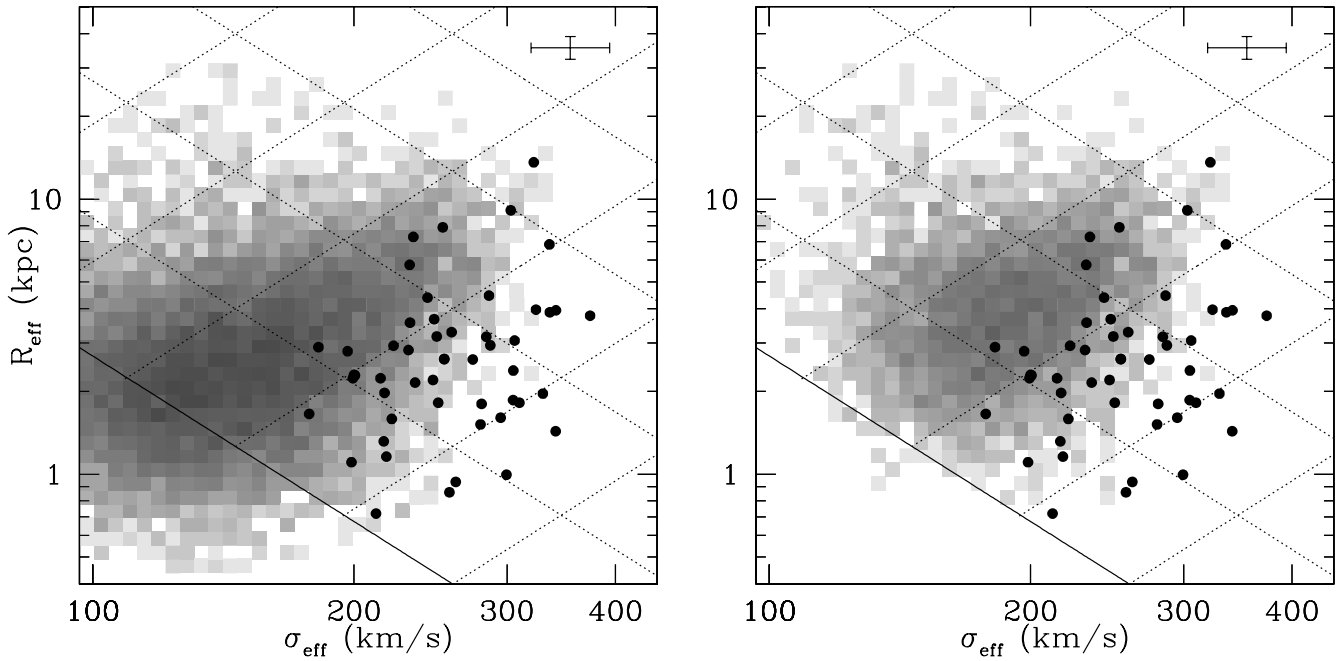

FIG. 3.—The $\sigma_{\rm eff}$-$R_{\rm eff}$ distributions of the nearby sample of early-type galaxies (*gray scale*) and the distant sample (*data points*). The red data points are cluster galaxies, and the blue data points are field galaxies. The error bars at the top right indicate the typical values of the errors for the distant sample. The solid line indicates $\log(M_{\rm dyn}) = 10.5$. Dotted lines indicate lines of constant $M_{\rm dyn}$ (*parallel to the solid line*), spaced by 0.5 dex, and lines of constant surface density $\Sigma_{\rm eff}$ (*perpendicular to the solid line*), also spaced by 0.5 dex. The left-hand panel shows the entire nearby sample; the right-hand panel only shows those galaxies in the nearby sample that would be included in the distant sample considering the selection effects that apply to the surveys (see § 5.1). The highly significant offset (>99.9% significance) between the local and distant samples implies significant structural evolution in the early-type galaxy population between $z \sim 1$ and the present. [*See the electronic edition of the Journal for a color version of this figure.*]

typical galaxies in the distant sample ($\sigma_{\rm eff} \sim 250$ km s^{-1}; $R_{\rm eff} \sim 3$ kpc) are rare in the local universe. In the nearby sample, galaxies with $\sigma_{\rm eff} \sim 250$ km s^{-1} have much larger sizes, and galaxies with sizes $R_{\rm eff} \sim 3$ kpc have dispersions of $\sigma_{\rm eff} \sim 150$ km s^{-1}. These numbers are only intended to guide the eye. A quantitative analysis of the differences between nearby and distant galaxies is presented below.

The distant sample is not directly comparable with the nearby sample in its entirety (Fig. 3, *left*), as the nearby sample reaches to much lower masses. In order to assess the question whether the true, underlying $\sigma_{\rm eff}$-$R_{\rm eff}$ distribution of distant galaxies is different from the $\sigma_{\rm eff}$-$R_{\rm eff}$ distribution of nearby galaxies, we need to remove the galaxies in the nearby sample that would not be included at $z \sim 1$ due to sample selection effects. The subsample of nearby galaxies that is observable at $z = 1$ is shown in the right panel of Figure 3. The two criteria that the galaxies in the observable subsample satisfy are $L > L_{\rm min}$ and $R_{\rm eff} < R_{\rm eff,\,max}$. $L_{\rm min} \sim 10^{10} L_{\odot,\,\rm B}$ is the luminosity limit for the field $z \sim 1$ surveys (see, e.g., van der Wel et al. 2005) after correcting for luminosity evolution between $z = 1$ and the present (0.555 dex; van Dokkum & van der Marel 2007). For the MS 1054−0321 cluster sample from Wuyts et al. (2004) this is $1.8 \times 10^{10} L_{\odot,\,\rm B}$. The second criterion $R_{\rm eff} < R_{\rm eff,\,max}$ takes into account that high signal-to-noise ratio (S/N) spectra are harder to obtain for low surface brightness galaxies than for high surface brightness galaxies with the same luminosity; i.e., the distant sample is biased in favor of small galaxies. The S/N of the spectra of van der Wel et al. (2005) and Treu et al. (2005b) do not precisely scale linearly with luminosity $L = {\rm IR}_{\rm eff}^2$, where I is the surface brightness, but as ${\rm S/N} \propto {\rm IR}_{\rm eff}^{1.6}$. This implies that, at fixed luminosity L, ${\rm S/N} \propto R^{-0.4}$. Since a dispersion measurement requires a minimum S/N ($\sim$12 Å^{-1}), a galaxy with luminosity L has a maximum radius $R_{\rm eff,\,max} \propto L^{2.5}$ for which its dispersion can be determined. We use the luminosity limits of the surveys discussed above to normalize the dependence between luminosity and maximum size; we simply assume that for the smallest galaxies ($R_{\rm eff} = 1$ kpc) the luminosity limit coincides with the size limit such that we have

$$R_{\rm eff,\,max}\ ({\rm kpc}) = \left(\frac{L}{L_{\rm min}}\right)^{2.5}. \qquad (5)$$

One would expect that for galaxies smaller than 3 kpc the signal-to-noise ratio of the spectra would not depend on size any longer since seeing generally dominates the apparent sizes of such small galaxies at $z \sim 1$. Because of the variety of telescopes, weather conditions, and data reduction techniques, this, however, is washed out and not apparent in the data. We note that the introduction of, effectively, a rudimentary surface brightness criterion is a step forward in modeling the selection effects with respect to earlier attempts that only take total luminosity into account.

The difference between the $\sigma_{\rm eff}$-$R_{\rm eff}$ distributions at low and high redshift is highly significant, even after taking selection effects into account (Fig. 3, *right*). The two-dimensional Kolgomorov-Smirnov statistic has a high value ($D = 3.71$), which implies that it is extremely unlikely that the nearby and distant samples are drawn from the same distribution. By repeatedly drawing samples from the nearby sample with the same size as the distant sample we confirm this: less than 0.001% of the simulated samples have $D = 3.71$ or higher.

5.2. *Evolution of the Mass-Radius Relation*

The structural difference between the nearby and distant samples described in the previous section implies that the $M_{\rm dyn}$-$R_{\rm eff}$ and $M_{\rm dyn}$-$\Sigma_{\rm eff}$ relations evolve with redshift. In Figure 4 we show the $M_{\rm dyn}$-$R_{\rm eff}$ relation for the distant sample and compare this with the equivalent relation for nearby galaxies derived in § 2.2. Clearly, the relation shifts to smaller radii from low to high redshift.

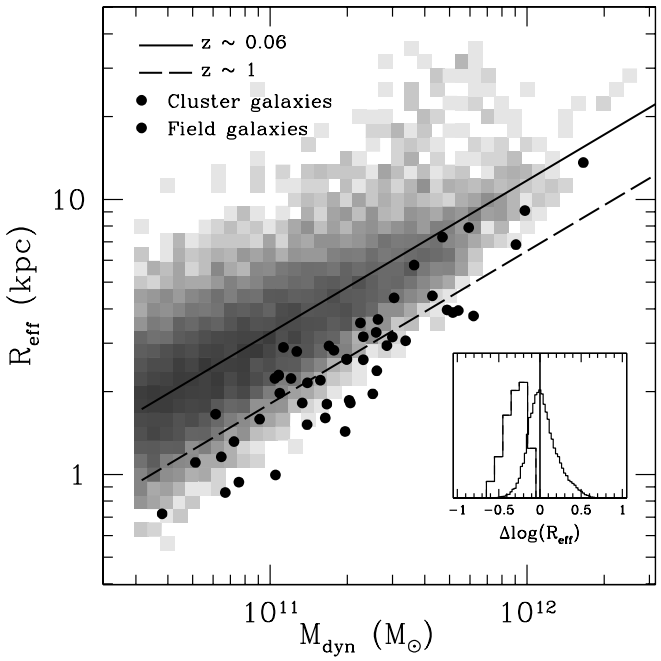

FIG. 4.—Mass-size relation for the nearby sample (*solid line*) and at $z \sim 1$ (*dashed line*); the symbols are the same as in Fig. 3. For the derivation of the $M_{\rm dyn}$-$R_{\rm eff}$ relation for the nearby sample see § 2.2; for the derivation of the $M_{\rm dyn}$-$R_{\rm eff}$ relation for the distant sample see § 5.2. The smaller, inset panel shows the distribution of the two samples around the $M_{\rm dyn}$-$R_{\rm eff}$ relation of the nearby sample (*the solid line in the large panel*). The distant galaxies are 1.8 ± 0.1 times smaller than the nearby galaxies. It appears that the most massive galaxies do not show as large an offset. This indicates that size evolution may be slower for the highest mass galaxies than for low-mass galaxies, but it has to be kept in mind that these very massive galaxies are brightest cluster galaxies and may therefore have developed differently from other galaxies. [*See the electronic edition of the Journal for a color version of this figure.*]

Parameterized as in equation (4) we find $R_c = 2.58 \pm 0.17$ and $b = 0.65 \pm 0.06$ with a scatter of 0.117 ± 0.013 dex (after subtracting the observational uncertainties in quadrature). The errors are estimated with a bootstrap/Monte Carlo simulation in which the data points are randomly sampled and varied according to the (correlated) measurement errors, which are assumed to be Gaussian. The systematic error in $R_{\rm eff}$ of 5% (see § 4) is also taken into account.

The treatment of the selection effects described in the previous section shows that the observed size evolution seen in Figure 4 is not an artifact. However, given the nature of the selection effects, which favor small galaxies over large galaxies, the intrinsic amount of size evolution and possible evolution in the slope and scatter of the $M_{\rm dyn}$-$R_{\rm eff}$ relation must be inferred through careful modeling. The goal is to derive the intrinsic $M_{\rm dyn}$-$R_{\rm eff}$ relation at $z \sim 1$ that reproduces the observed $M_{\rm dyn}$-$R_{\rm eff}$ distribution after applying the selection criteria. We take an iterative approach due to the interdependence of the selection criteria and the amount of evolution in zero point, slope, and scatter of the $M_{\rm dyn}$-$R_{\rm eff}$ relation. In the following we de-evolve the properties of the nearby sample to constrain the form of the true, underlying $z \sim 1$ $M_{\rm dyn}$-$R_{\rm eff}$ relation.

The simplest evolutionary scenario is a change in the zero point R_c (see eq. [4]). For each object in the nearby sample the size is reduced by the same amount $\Delta \log (R_{\rm eff})$, and those that do not satisfy the selection criteria described in the previous section are

removed. From the remaining subsample the "observed" $M_{\rm dyn}$-$R_{\rm eff}$ relation is determined. The different selection criteria and sample sizes for field and cluster galaxies are taken into account in this process, which is repeated for many different values of $\Delta \log (R_{\rm eff})$. We find that an intrinsic value of $R_c = 2.64 \pm 0.18$ reproduces the observed value of $R_c = 2.58 \pm 0.17$. Hence, it appears that selection effects do not strongly affect the inferred size evolution.

However, the scatter of the assumed intrinsic distribution (0.14 dex) is higher than the observed scatter (0.117 ± 0.013 dex). This cannot be explained by selection effects in the simple scenario described above. It is therefore required that the scatter, as well as the zero point, is also treated as an evolving parameter. This is implemented in our analysis by reducing or increasing the offset of each galaxy in the nearby sample from the best-fit $M_{\rm dyn}$-$R_{\rm eff}$ relation by a given fraction. Doing so, we find that the best-fitting zero point R_c is not different from the earlier estimate based on a nonevolving scatter. We also find that the evidence for evolution in the scatter is weak ($\sim$1.5 σ). This exercise mainly serves to show that our size-evolution result in not sensitive to the amount of evolution in the scatter allowed by the observations.

A similar verification must be carried out for evolution in the slope of the $M_{\rm dyn}$-$R_{\rm eff}$ relation. Allowing only the scatter and the size to evolve, as described above, the inferred slope of the "observed" $M_{\rm dyn}$-$R_{\rm eff}$ relation is 0.59, marginally consistent with the true observed slope of $b = 0.65 \pm 0.06$. If we treat the slope as an additional, third free parameter we confirm that evolution in the slope, as constrained by our measurements, does not affect our size-evolution measurement. An intrinsic slope of $b = 0.61$ provides a better fit than the original slope of the $M_{\rm dyn}$-$R_{\rm eff}$ relation of the nearby sample ($b = 0.56$), but the difference is marginal ($\sim$1 σ).

We conclude that, despite (weak) evidence for evolution in the slope and the scatter of the $M_{\rm dyn}$-$R_{\rm eff}$ relation with redshift, there is no significant improvement in modeling the observations by adopting slope and scatter as free parameters. Modeling the evolution by a fractional change in size, regardless of mass and offset from the local $M_{\rm dyn}$-$R_{\rm eff}$ relation, provides an equally good fit. Most importantly, changing the slope and scatter within the range allowed by the observations does not affect the inferred size evolution. We find that $R_c = 2.64 \pm 0.18$ kpc at $z = 0.90$, a factor of 1.8 ± 0.1 times smaller than at $z \sim 0.06$.

The weak evidence for a change in slope of the $M_{\rm dyn}$-$R_{\rm eff}$ relation may also be interpreted as a difference between field and cluster galaxies, as the more massive galaxies in our sample tend to be cluster galaxies. Assuming that slope and scatter remain constant but that the zero point of the $M_{\rm dyn}$-$R_{\rm eff}$ relation evolves differently for field and cluster galaxies, we find that $R_c = 2.49$ kpc for field galaxies and $R_c = 3.06$ kpc for cluster galaxies. The 1 σ error on this difference of 0.57 kpc is 0.32 kpc. The true error may be larger since in this estimate it is assumed that scatter and slope behave the same in the different environments and that there are no relative systematic errors in the size determinations of field and cluster galaxies. The evidence for a difference between the size evolution of field and cluster early-type galaxies is therefore weak (see also Rettura et al. 2008). However, we have to keep in mind that so far only a very small number of clusters is considered. Future studies will need to extend the existing analyses to a larger number of clusters to verify the general validity of the results.

So far, we have assumed that the masses of the galaxies do not change. Our justification is that the scatter hardly depends on mass; the effect of a changing mass function on modeling selection effects is expected to be small. However, physically speaking, it is unnatural to propose size evolution without changes in the masses of galaxies. Moreover, if the characteristic mass above which the number density of galaxies drops off exponentially evolves with redshift, selection effects will change as well. The simplest way to implement mass evolution is to assume that $M \propto R_{\rm eff}$ (eq. [2]). Including this in our modeling procedure shows that the effect on the inferred size evolution is less than 5%, and we therefore adopt the results with no mass evolution.

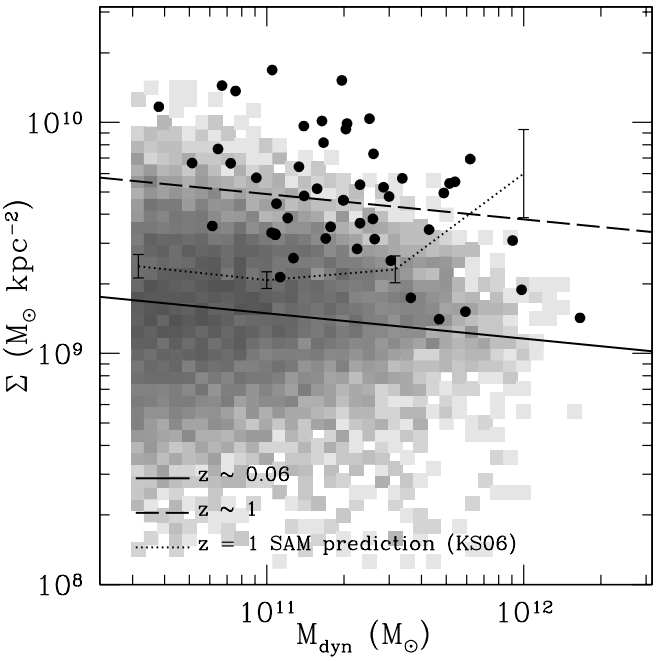

FIG. 5.—Mass-density relation at $z \sim 1$. The symbols and lines are the same as in Figs. 3 and 4. The $z \sim 1$ early-type galaxies are $\sim$4 times more dense than their nearby counterparts. The prediction of the semianalytic size-evolution model for elliptical galaxies from Khochfar & Silk (2006a) is shown as the dotted line. The error bars indicate the predicted size evolution between $z = 0.8$ and $z = 1.2$, the redshift range of our distant sample. Despite qualitative agreement, there are significant quantitative differences between the predicted and observed evolution. [*See the electronic edition of the Journal for a color version of this figure.*]

We recall that the nearby sample is biased against compact early-type galaxies (§ 2.1). The potentially underestimated number of galaxies with dispersions $\sigma > 300$ km s^{-1} is unlikely to drastically affect the size-evolution determination for the sample as a whole as the average dispersion of the galaxies in the distant sample is smaller than that. However, the slope of the local $M_{\rm dyn}$-$R_{\rm eff}$ relation is possibly overestimated, which would lead to an underestimate of the slope evolution. More important is the problem that small galaxies are missed because of their photometric misclassification as stars in the SDSS. To fully address this issue a complete analysis of the SDSS photometric catalog is required, which is clearly beyond the scope of this paper. However, we can say that it is highly unlikely that the average size of nearby early-type galaxies is underestimated by a factor of 2 because of this bias. On the other hand, for the interpretation of our results and identifying the mechanisms responsible for size evolution (see § 6.3) this bias could prove to be important.

In § 3 we noted that the dynamical mass estimate as adopted in this paper (eq. [2], with $\beta = 5$) may be too low for rotating early-type galaxies. If this is the case, then size evolution for these galaxies will be underestimated by $\sim$10%. Since this is within the uncertainties of our measurements we do not take this further into account.

Obviously, size evolution at fixed mass translates into density evolution. This is illustrated in Figure 5 where we compare the density distribution of $z \sim 1$ early-type galaxies with the $M_{\rm dyn}$-$\Sigma_{\rm eff}$ relation for nearby galaxies. Because $\Sigma_{\rm eff}$ does not strongly depend

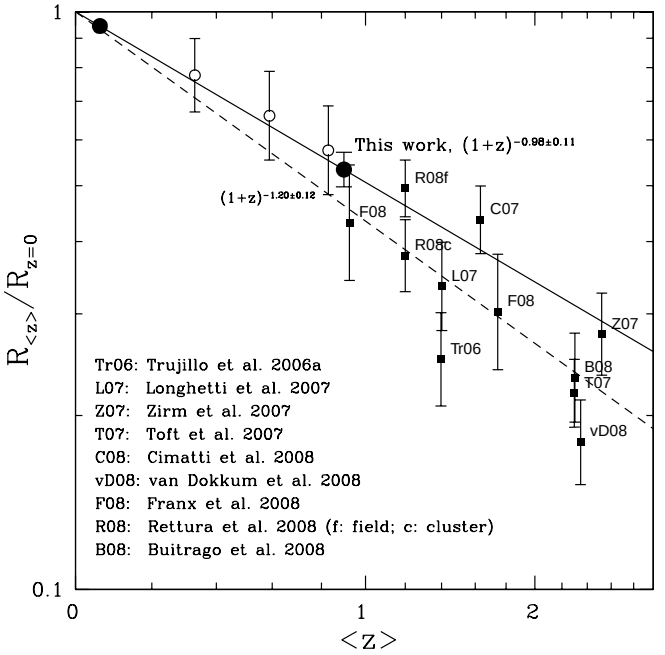

FIG. 6.— Size evolution with redshift as derived in this paper with dynamically determined masses (*large filled circles*) compared with previous results based on photometric masses (*small filled squares*). The solid line connects our samples at $z \sim 0.06$ and $z \sim 1$; the dashed line is a linear least-squares fit to the small filled data points. The open circles are samples of cluster galaxies with photometrically measured masses and serve as an illustration that size evolution shows a continuous trend between $z = 2.5$ and the present. The broad agreement in size evolution as derived from galaxies with dynamically and photometrically determines masses reinforces the conclusions of previous, photometric studies whose results were potentially mitigated by considerable systematic effects that do not affect our analysis.

on $M_{\rm dyn}$, evolutionary trends are readily visible; $z \sim 1$ early-type galaxies are $\sim$4 times more dense than their local counterparts. The apparent change in slope can possibly be explained by selection effects, completely analogous to our conclusion that this is the case with the $M_{\rm dyn}$-$R_{\rm eff}$ relation. Note that compared to the increase in projected density, the increase in physical density will be even larger.

Until recently, early-type galaxies were thought to evolve more or less passively. This appears to be an oversimplification and may apply more to their stellar populations than to their structural properties. In the following section we discuss possible explanations in the context of theoretical predictions and the comparison with results from studies with different observational strategies.

6. DISCUSSION

6.1. *Comparison with Photometric Size-Evolution Measurements*

The main goal of this paper is to use dynamical measurements to investigate whether early-type galaxies were smaller and denser in the past. Previous work has shown that the stellar mass surface density is higher, but there are a number of issues with such studies as they rely on stellar population models and they ignore possible changes in the underlying dark matter profile.

In Figure 6 we compare the size-evolution results presented in the previous section with size-evolution results for early-type galaxies based on photometric mass estimates. For all the literature samples we take the mean redshift and the mean stellar mass (normalized to the Kroupa IMF) and compute the mean offset from the local mass-size relation from Shen et al. (2003). We include four intermediate-redshift cluster galaxy samples with photometrically measured masses and sizes from WFPC2 or ACS imaging. The data are described by Holden et al. (2007), and the sizes are measured as described in this paper (§ 3). These four clusters are CL 1358+62 at $z = 0.33$, MS 2053−04 at $z = 0.59$, and MS 1054−0321 and RX J0152.7−1357, both at $z = 0.83$. Note that we also include MS 1054−0321 in the present study with dynamical mass measurements. The agreement between the independent measurements confirms that at least out to $z \sim 1$ dynamical and photometric mass estimates based on optical colors and spectral energy distributions agree within the statistical errors as was previously shown by van der Wel et al. (2006), Rettura et al. (2006), and Holden et al. (2006).

The literature samples have all been selected in different ways, and so a direct comparison with our work may not be straightforward. Not all samples are morphologically selected; many are selected by their spectral or photometric properties. In the local universe there is substantial overlap between samples of early-type galaxies that are selected by different criteria; therefore, it is a reasonable assumption to suppose this to also be the case at high redshift, where different indicators (low star formation rates, red colors, smooth visual appearance) also reflect a common nature. Recently, several studies have shown hints that this is indeed the case (e.g., van Dokkum et al. 2008; Kriek et al. 2008), but these issues need to be further addressed in the future.

Even with this cautionary proviso, the broad agreement between the results presented in this paper and the photometric results at higher redshifts is striking. All studies included in Figure 6 are consistent with significant size evolution of several factors between $z \sim 1$–2 and the present for galaxies with a given mass. A linear fit in log-log space to our two data points at $z \sim 0.06$ and $z \sim 1$ gives $R_{\rm eff}(z) \propto (1+z)^{-0.98\pm0.11}$. With a linear fit to the photometric data the inferred rate of evolution is $R_{\rm eff}(z) \propto (1+z)^{-1.20\pm0.12}$, where the error is obtained via a bootstrap/Monte Carlo simulation.

The broad agreement of our measurement of the size evolution of early-type galaxies with the photometric studies is encouraging and alleviates concerns about serious systematic effects that potentially could have compromised previous work. Most notably, uncertainties in the photometric mass estimates used in all other previous work appear to have a limited impact, at least compared to factors of $\gtrsim$5 which would mimic the strong, observed size and density evolution. Uncertainties in photometric mass estimates on the level of a factor of $\sim$2 due to differences among the various stellar population models (e.g., Bruzual & Charlot 2003; Maraston 2005) remain an issue, but to invoke, for example, an unconventional stellar IMF as an alternative to radically different structural properties of high-redshift early-type galaxies is no longer necessary.

Other systematic uncertainties cannot explain the observed evolution either. In our size measurements, systematic effects have been taken into account (see §§ 3.2 and 4). We are confident, for example, that we would detect low-surface brightness envelopes around distant galaxies. Furthermore, we know that only a minority of morphologically selected $z \sim 1$ early-type galaxies ($\sim$10%) show signs of nuclear activity (e.g., Rodighiero et al. 2007; van der Wel et al. 2007) such that it is unlikely that central point sources affect our size measurements. This is also clear from the fact that the residuals of our $R^{1/4}$ profile fits generally do not show central point sources and that none of the deep spectra used to measure dispersions show evidence for nuclear activity. Furthermore, the good correspondence between the rest-frame wavelength of the imaging data sets used at different redshifts assures us that morphological K-corrections do not play a significant role.

Despite the broad consistency between our results and those previously published, the agreement is not perfect. There is a marginal

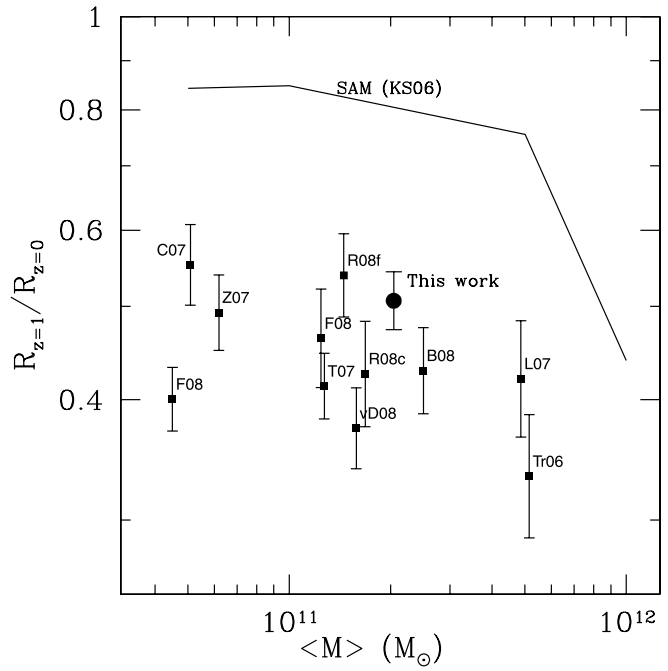

FIG. 7.—Size evolution per unit redshift vs. mean galaxy mass of our sample (*large circle*) and samples taken from the literature (*small squares*; see Fig. 6 for references). Samples consisting of high-mass galaxies show somewhat stronger size evolution than samples consisting of low-mass galaxies, which is qualitatively, but not quantitatively, consistent with the predictions from the semianalytic model from Khochfar & Silk (2006a) (*solid line*). This conclusion should be considered highly tentative, however, as this interpretation is hampered by systematic uncertainties and small sample sizes.

inconsistency at the 1.5 σ level between the size-evolution measurement from kinematic data and the size-evolution measurement from photometric data shown in Figure 6. This could point to the presence of some systematic effects within the $z > 1.5$ results. Alternatively, the different studies sample galaxies with a wide range in masses, and therefore mass-dependent size evolution could lead to apparent discrepancies among the samples. This is explored in the following section.

Our robust results strengthen the results from previous studies. We conclude that early-type galaxies at $z = 1$ are $\sim$2 times smaller than local early types with the same mass, and that at $z =$ 2–2.5 this size difference is likely increased to a factor of $\sim$4, as previously observed by Zirm et al. (2007), Toft et al. (2007), van Dokkum et al. (2008), and Buitrago et al. (2008).

6.2. *Comparison with Model Predictions*

The fact that we see considerable evolution in galaxy size with redshift is not surprising from a theoretical perspective. Most semianalytic models of galaxy formation in a ΛCDM universe predict substantial size evolution over the past several billion years. A comparison between the observed and model-predicted amount of size evolution will help to identify the mechanism(s) that are responsible. In Figure 5 we directly compare the observed evolution in surface density with the predictions from the semianalytic work by Khochfar & Silk (2006a). For galaxies with a given mass the model significantly underpredicts the evolution in size and density, except, perhaps, for the most massive galaxies. In our data set we see no indication that the magnitude of size and density evolution increases with galaxy mass, as predicted by the models. In fact, the most massive galaxies in our sample are precisely the only ones that are not different from local massive galaxies. Note, however, that statistically speaking the evidence for mass-dependent evolution is weak (see § 5.2). Moreover, the most massive galaxies in our distant sample are a special subset, brightest cluster galaxies. Such galaxies have been shown to have properties that deviate from those of other massive galaxies (see, e.g., von der Linden et al. 2007; Bernardi et al. 2007).

By including the $z = 1.5$–2.5 photometric samples discussed in § 6.1 we can place further constraints on the models. In Figure 7 we compare the observed size evolution of the available samples, normalized to $z = 1$, with the model predictions from Khochfar & Silk (2006a). Representing the model predictions by a single line is justified by the fact that the predicted evolution of $\log(R_{\rm eff})$ with $\log(1+z)$ is very close to linear. Again, the observed size evolution is stronger than that predicted by the model.

It is interesting to note that, in qualitative agreement with the model prediction, we see a hint that size evolution depends on mass in the compilation presented in Figure 7. The samples containing, on average, the lowest mass galaxies display marginally less evolution. It has to be kept in mind, however, that small sample sizes and systematic effects are more important for determining second-order effects such as mass dependence (see § 5.2). A clue that systematic uncertainties may play a role is the remaining difference between the kinematic and photometric samples. Alternatively, it may signify nonlinear evolution of $\log(R_{\rm eff})$ with $\log(1+z)$.

6.3. *Size Evolution of Individual Galaxies*

It appears that the observed size evolution of a factor of $\sim$2 between $z = 1$ and the present for early-type galaxies with masses $\sim 10^{11}\,M_\odot$ is similar to the predicted evolution for early-type galaxies that are an order of magnitude more massive (see Figs. 5 and 7). This suggests that the mechanism responsible for increasing the average size of early-type galaxies with time may be well understood, but that it is not implemented correctly in the current semianalytic model from Khochfar & Silk (2006a). The process of size evolution may occur at different times and under different circumstances than is now assumed. This may be related to the late assembly of very massive galaxies in models of this kind (see also, e.g., De Lucia et al. 2006), a prediction that is challenged by various observations (e.g., Cimatti et al. 2006; Scarlata et al. 2007; Cool et al. 2008).

It is beyond the scope of this paper to fully discuss these possible discrepancies. Instead we will explore the question whether the proposed physical processes responsible for size evolution are consistent with the observed trends. In the semianalytic models it is assumed (and this is confirmed by numerical simulations) that mergers drive size evolution. The gas content of merging galaxies largely determines the relative size of the merger remnant compared to its ancestors. Because gas fractions were higher in the past, galaxies that form early will be smaller than galaxies that form late. In the framework of cosmological simulations this means that galaxies at high redshift will be smaller because they were formed through gas-rich mergers and that those merger remnants can grow over time through subsequent mergers with other galaxies that are progressively more devoid of gas.

The question is whether the observed size evolution is dominated by size evolution of individual galaxies or simply by the addition of larger galaxies over time. At $z = 1$ only about 30%–50% of the present-day early-type galaxy evolution had formed (Bell et al. 2004; Brown et al. 2007; Scarlata et al. 2007; Faber et al. 2007). If we assume that these galaxies will make up the 30%–50% most dense early-type galaxies in the present-day universe, then the scatter in the local $M_{\rm dyn}$-$R_{\rm eff}$ relation implies that the sample-averaged size increases by a factor of 1.3–1.4 between $z = 1$ and the present. Such evolution is thus expected in the

absence of size evolution of individual galaxies, and this is less than the observed evolution of a factor of 2. To explain the observed evolution by growth of the early-type galaxy population without changes in the sizes of individual galaxies, the number density of early-type galaxies is required to increase by an order of magnitude between $z = 1$ and the present. Such strong evolution is clearly ruled out by the above-mentioned determinations of the number density of red galaxies at $z \sim 1$.

Similarly, at $z = 2$ only $\sim$10% of the galaxies with masses $\gtrsim 10^{11}\ M_{\odot}$ had been assembled (Kriek et al. 2008); if those galaxies evolve into the 10% most dense present-day early-type galaxies, then an increase in average size by a factor of $\sim$2 can be accounted for, less than the observed amount of evolution. These arguments are in agreement with the conclusions from Cimatti et al. (2008), who show that local galaxies with the same sizes and masses as galaxies in the $z = 1$–2 samples are so rare in the local universe that it can be confidently ruled out that their structure remains unchanged up until the present day. Note, however, that these arguments may be affected by the aforementioned biases in the SDSS (§ 2.1).

We conclude that size evolution due to the addition of larger galaxies over time contributes at most half of the observed evolution in the $M_{\rm dyn}$-$R_{\rm eff}$ relation. The remainder must be due to size evolution of individual galaxies. Numerical simulations have demonstrated that when early-type galaxies accrete neighbors without significant dissipational processes, $\sigma_{\rm eff}$ does not change by much and that, to first order, $R_{\rm eff}$ increases linearly with mass. This does not depend strongly on the mass of the accreted object, i.e., the mass ratio of the merger (Boylan-Kolchin et al. 2005, 2006; Robertson et al. 2006).

Simulations in a cosmological context show that an increase in size by a factor of 2 between $z \sim 1$ and the present is certainly possible (Naab et al. 2007). The strong observed size evolution thus argues in favor of a scenario in which significant mass from low-mass companions is accreted onto existing early-type galaxies over the past $\sim$7 Gyr, which also explains the broad tidal features that are frequently observed around early-type galaxies (van Dokkum 2005). As shown by Feldmann et al. (2008) such features are not necessarily, and are even quite unlikely to be, the result of major merger events and are most likely due to the accretion of low-mass, gas-poor satellites.

We note that the size evolution of individual galaxies and the evolution of the sample average are inseparable because galaxies evolve in mass as well as in size. Nonetheless, it is important to distinguish this complex scenario from the simple picture in which early-type galaxies that form at different redshifts have different sizes but do not structurally evolve at later times. The strong observed size evolution clearly rules out the latter, indicating that the buildup of the early-type galaxy population is a complex and ongoing process.

Finally, it is remarkable that the change in the sizes of early-type galaxies is consistent with and differs by less than 15% from the change in the scale factor of the universe, $1 + z$. Within the standard cold dark matter scenario this is likely a coincidence since dissipational, strongly nonlinear processes that are decoupled from cosmic expansion dominate at the kiloparsec scale of forming galaxies. Nonetheless, we cannot exclude the possibility that there is an underlying, fundamental reason that galaxies are scale-invariant with respect to a comoving coordinate system. In an alternative description of dark matter, i.e., Bose-Einstein condensed, ultralight particles with a $\sim$10 kpc–sized wave function (fuzzy dark matter or FDM; Sin 1994; Hu et al. 2000), sizes of halos and their occupying galaxies possibly follow the cosmic expansion rate (Lee 2008).

7. SUMMARY

In § 2 we construct a large sample of nearby ($0.04 < z < 0.08$) early-type galaxies extracted from the SDSS (DR6). We use the pipeline velocity dispersion measurements and obtain our own size measurements in order to construct the local dynamical mass-size relation (§ 2.2). In addition, we construct a sample of 50 morphologically selected early-type galaxies in the redshift range $0.8 < z < 1.2$ with measured velocity dispersions (§ 3). Sizes are determined from ACS imaging in the same manner as for the galaxies in the nearby sample, and systematic effects are quantified through simulations (§ 4). The distant sample contains galaxies in the mass range $3 \times 10^{10}\ M_{\odot} < M \lesssim 10^{12}\ M_{\odot}$, with a typical mass of $2 \times 10^{11}\ M_{\odot}$.

The main result is that the $\sigma_{\rm eff}$-$R_{\rm eff}$ distributions of the nearby and distant samples are significantly different, even after we correct for the incompleteness of the distant sample at low masses (§ 5). The implied size evolution is $R_{\rm eff} \propto (1 + z)^{-0.98 \pm 0.11}$, or a factor of 1.97 ± 0.15 between $z = 1$ and the present. Similarly, the projected surface densities of the distant early-type galaxies are a factor of $\sim$4 higher than those of their local counterparts. The stellar populations of early-type galaxies that already existed at $z = 1$ may, for the most part, be passively evolving over the past 7–8 Gyr; however, their structural properties undergo substantial changes over that period.

Our results are in broad agreement (see § 6.1) with previously published size-evolution measurements that are based on samples without dynamical mass measurements and, in some cases, without spectroscopic redshifts, high-resolution *HST* imaging, and/or consistently determined sizes. We therefore conclude that systematic effects, most notably those in the mass estimates, which potentially could have hampered previous studies are small relative to the observed amount of evolution.

The observed size evolution is in qualitative agreement with predictions from recent semianalytic models. However, the predicted evolution is much slower than the observed evolution. The observed size evolution of early-type galaxies can be understood within the context of the cold dark matter scenario in which galaxies that form late have larger sizes than galaxies that form early, due to lower gas fractions at late times, and the growth of individual galaxies through the mostly dissipationless accretion of satellites at later evolutionary stages.

We thank the referee for his helpful review. A. V. D. W. would like to thank Jenny Graves for sharing her SDSS catalog of early-type galaxies, Eric Bell, Hans-Walter Rix, and Pieter van Dokkum for interesting discussions, and Sadegh Khochfar for providing his model predictions. Support from NASA grant NAG5-7697 is also gratefully acknowledged. The authors wish to recognize and acknowledge the very significant cultural role and reverence that the summit of Mauna Kea has always had within the indigenous Hawaiian community. We are most fortunate to have the opportunity to conduct observations from this mountain. Funding for the SDSS and SDSS-II has been provided by the Alfred P. Sloan Foundation, the Participating Institutions, the National Science Foundation, the US Department of Energy, the National Aeronautics and Space Administration, the Japanese Monbukagakusho, the Max Planck Society, and the Higher Education Funding Council for England. The SDSS Web Site is http://www.sdss.org/. The SDSS is managed by the Astrophysical Research Consortium for the Participating Institutions. The Participating Institutions are the American Museum of Natural History, Astrophysical Institute Potsdam, University of

Basel, University of Cambridge, Case Western Reserve University, University of Chicago, Drexel University, Fermilab, the Institute for Advanced Study, the Japan Participation Group, Johns Hopkins University, the Joint Institute for Nuclear Astrophysics, the Kavli Institute for Particle Astrophysics and Cosmology, the Korean Scientist Group, the Chinese Academy of Sciences (LAMOST), Los Alamos National Laboratory, the Max-Planck-Institute for Astronomy (MPIA), the Max-Planck-Institute for Astrophysics (MPA), New Mexico State University, Ohio State University, University of Pittsburgh, University of Portsmouth, Princeton University, the United States Naval Observatory, and the University of Washington.

REFERENCES

Adelman-McCarthy, J. K., et al. 2008, ApJS, 175, 297
Bell, E. F., et al. 2004, ApJ, 608, 752
Bernardi, M. 2007, AJ, 133, 1954
Bernardi, M., Hyde, J. B., Sheth, R. K., Miller, C. J., & Nichol, R. C. 2007, AJ, 133, 1741
Blakeslee, J. P., et al. 2006, ApJ, 644, 30
Boylan-Kolchin, M., Ma, C.-P., & Quataert, E. 2005, MNRAS, 362, 184
———. 2006, MNRAS, 369, 1081
Brown, M. J. I., et al. 2007, ApJ, 654, 858
Bruzual, G., & Charlot, S. 2003, MNRAS, 344, 1000
Buitrago, F., Trujillo, I., Conselice, C. J., Bouwens, R. J., Dickinson, M., & Yan, H. 2008, ApJ, submitted (arXiv:0807.4141)
Cappellari, M., et al. 2006, MNRAS, 366, 1126
Chabrier, G. 2003, PASP, 115, 763
Cimatti, A., Daddi, E., & Renzini, A. 2006, A&A, 453, L29
Cimatti, A., et al. 2008, A&A, 482, 21
Cool, R. J., et al. 2008, ApJ, 682, 919
Daddi, E., et al. 2005, ApJ, 626, 680
Davé, R. 2008, MNRAS, 385, 147
De Lucia, G., Springel, V., White, S. D. M., Croton, D., & Kauffmann, G. 2006, MNRAS, 366, 499
de Vaucouleurs, G. 1948, Ann. d'Astrophys., 11, 247
di Serego Alighieri, S., et al. 2005, A&A, 442, 125
Djorgovski, S., & Davis, M. 1987, ApJ, 313, 59
Dressler, A., Lynden-Bell, D., Burstein, D., Davies, R. L., Faber, S. M., Terlevich, R., & Wegner, G. 1987, ApJ, 313, 42
Faber, S. M., et al. 2007, ApJ, 665, 265
Fardal, M. A., Katz, N., Weinberg, D. H., & Davé, R. 2007, MNRAS, 379, 985
Feldmann, R., Mayer, L., & Carollo, C. M. 2008, ApJ, 684, 1062
Ford, H. C., et al. 1998, Proc. SPIE, 3356, 234
Franx, M., van Dokkum, P. G., Foerster Schreiber, N. M., Wuyts, S., Labbe, I., & Toft, S. 2008, ApJ, submitted (arXiv:0808.2642)
Giavalisco, M., et al. 2004, ApJ, 600, L93
Graves, G. J., Faber, S. M., Schiavon, R. P., & Yan, R. 2007, ApJ, 671, 243
Holden, B. P., et al. 2005, ApJ, 620, L83
———. 2006, ApJ, 642, L123
———. 2007, ApJ, 670, 190
Hoversten, E. A., & Glazebrook, K. 2008, ApJ, 675, 163
Hu, W., Barkana, R., & Gruzinov, A. 2000, Phys. Rev. Lett., 85, 1158
Jørgensen, I., Bergmann, M., Davies, R., Barr, J., Takamiya, M., & Crampton, D. 2005, AJ, 129, 1249
Jørgensen, I., Franx, M., & Kjaergaard, P. 1995, MNRAS, 276, 1341
Khochfar, S., & Silk, J. 2006a, ApJ, 648, L21
———. 2006b, MNRAS, 370, 902
Kriek, M., van der Wel, A., van Dokkum, P. G., Franx, M., & Illingworth, G. D. 2008, ApJ, 682, 896
Kriek, M., et al. 2006, ApJ, 649, L71
Krist, J. 1995, in ASP Conf. Ser. 77, Astronomical Data Analysis Software and Systems IV, ed. R. A. Shaw, H. E. Payne, & J. J. E. Hayes (San Francisco: ASP), 349
Kroupa, P. 2001, MNRAS, 322, 231
Larson, R. B. 2005, MNRAS, 359, 211
Lee, J.-W. 2008, preprint (arXiv:0805.2877)
Loeb, A., & Peebles, P. J. E. 2003, ApJ, 589, 29
Longhetti, M., et al. 2007, MNRAS, 374, 614
Maraston, C. 2005, MNRAS, 362, 799
Mei, S., et al. 2007, ApJ, 655, 144
Naab, T., Johansson, P. H., Ostriker, J. P., & Efstathiou, G. 2007, ApJ, 658, 710
Paturel, G., et al. 2003, A&A, 412, 45
Peng, C. Y., Ho, L. C., Impey, C. D., & Rix, H.-W. 2002, AJ, 124, 266
Rettura, A., et al. 2006, A&A, 458, 717
———. 2008, A&A, in press (arXiv:0806.4604)
Robertson, B., Cox, T. J., Hernquist, L., Franx, M., Hopkins, P. F., Martini, P., & Springel, V. 2006, ApJ, 641, 21
Rodighiero, G., et al. 2007, MNRAS, 376, 416
Salpeter, E. E. 1955, ApJ, 121, 161
Scalo, J. M. 1986, Fundam. Cosm. Phys., 11, 1
Scarlata, C., et al. 2007, ApJS, 172, 494
Sérsic, J. L. 1968, Atlas de Galaxias Australes (Cordoba: Obs. Astron.)
Shen, S., et al. 2003, MNRAS, 343, 978
Sin, S.-J. 1994, Phys. Rev. D, 50, 3650
Toft, S., et al. 2007, ApJ, 671, 285
Treu, T., Ellis, R. S., Liao, T. X., & van Dokkum, P. G. 2005a, ApJ, 622, L5
Treu, T., et al. 2005b, ApJ, 633, 174
Trujillo, I., Conselice, C. J., Bundy, K., Cooper, M. C., Eisenhardt, P., & Ellis, R. S. 2007, MNRAS, 382, 109
Trujillo, I., et al. 2004, ApJ, 604, 521
———. 2006a, MNRAS, 373, L36
———. 2006b, ApJ, 650, 18
van der Marel, R. P., & van Dokkum, P. G. 2007, ApJ, 668, 756
van der Wel, A., Franx, M., Illingworth, G. D., & van Dokkum, P. G. 2007, ApJ, 666, 863
van der Wel, A., Franx, M., van Dokkum, P. G., & Rix, H.-W. 2004, ApJ, 601, L5
van der Wel, A., Franx, M., van Dokkum, P. G., Rix, H.-W., Illingworth, G. D., & Rosati, P. 2005, ApJ, 631, 145
van der Wel, A., Franx, M., Wuyts, S., van Dokkum, P. G., Huang, J., Rix, H.-W., & Illingworth, G. 2006, ApJ, 652, 97
van der Wel, A., & van der Marel, P. M. 2008, ApJ, 684, 260
van Dokkum, P. G. 2005, AJ, 130, 2647
———. 2008, ApJ, 674, 29
van Dokkum, P. G., & Ellis, R. S. 2003, ApJ, 592, L53
van Dokkum, P. G., Franx, M., Kelson, D. D., & Illingworth, G. D. 1998, ApJ, 504, L17
van Dokkum, P. G., & Stanford, S. A. 2003, ApJ, 585, 78
van Dokkum, P. G., & van der Marel, R. P. 2007, ApJ, 655, 30
van Dokkum, P. G., et al. 2008, ApJ, 677, L5
von der Linden, A., Best, P. N., Kauffmann, G., & White, S. D. M. 2007, MNRAS, 379, 867
Wuyts, S., van Dokkum, P. G., Kelson, D. D., Franx, M., & Illingworth, G. D. 2004, ApJ, 605, 677
Yan, R., Newman, J. A., Faber, S. M., Konidaris, N., Koo, D., & Davis, M. 2006, ApJ, 648, 281
York, D. G., et al. 2000, AJ, 120, 1579
Zirm, A. W., et al. 2007, ApJ, 656, 66