The Astrophysical Journal, 642:653–672, 2006 May 10

KECK DEEP FIELDS. II. THE ULTRAVIOLET GALAXY LUMINOSITY FUNCTION AT $z \sim 4$, 3, AND 2[1]

Marcin Sawicki[2]
Dominion Astrophysical Observatory, Herzberg Institute of Astrophysics, National Research Council, 5071 West Saanich Road, Victoria, BC V9E 2E7, Canada; and Caltech Optical Observatories, California Institute of Technology, MS 320-47, Pasadena, CA 91125; marcin.sawicki@nrc.gc.ca

and

David Thompson
Caltech Optical Observatories, California Institute of Technology, MS 320-47, Pasadena, CA 91125; djt@irastro.caltech.edu
Received 2005 April 25; accepted 2005 December 14

ABSTRACT

We use very deep U_nGRI multifield imaging obtained at the Keck telescope to study the evolution of the rest-frame 1700 Å galaxy luminosity function as the universe doubles its age from $z \sim 4$ to ~ 2. We use exactly the same filters and color-color selection as those used by the Steidel team but probe significantly fainter limits, well below L^*. The depth of our imaging allows us to constrain the faint end of the luminosity function, reaching $M_{1700} \sim -18.5$ at $z \sim 3$ (equivalent to $\sim 1\ M_\odot\ \mathrm{yr}^{-1}$), accounting for both $N^{1/2}$ uncertainty in the number of galaxies and cosmic variance. We carefully examine many potential sources of systematic bias in our LF measurements before drawing the following conclusions. We find that the luminosity function of Lyman break galaxies evolves with time and that this evolution is differential with luminosity. The result is best constrained between the epochs at $z \sim 4$ and ~ 3, where we find that the number density of sub-L^* galaxies increases with time by at least a factor of 2.3 (11 σ statistical confidence); while the faint end of the LF evolves, the bright end appears to remain virtually unchanged, indicating that there may be differential, luminosity-dependent evolution (98.5% statistical probability). Potential systematic biases restrict our ability to draw strong conclusions about continued evolution of the luminosity function to lower redshifts, $z \sim 2.2$ and ~ 1.7, but, nevertheless, it appears certain that the number density of $z \sim 2.2$ galaxies at all luminosities we studied, $-22 > M_{1700} > -18$, is at least as high as that of their counterparts at $z \sim 3$. While it is not yet clear what mechanism underlies the observed evolution, the fact that this evolution is differential with luminosity opens up new avenues of improving our understanding of how galaxies form and evolve at high redshift.

Subject headings: galaxies: evolution — galaxies: formation — galaxies: high-redshift — galaxies: starburst

1. INTRODUCTION

Understanding the formation and evolution of galaxies continues to be one of the most active fields of observational cosmology. Over the last decade, advances in instrumentation and technique have made it possible to select large samples of "normal" star-forming galaxies for direct study at redshifts that correspond to a time when the universe was only a tenth of its present age and so study galaxy assembly at a time when galaxies were young. Several different approaches for selecting high-z galaxies are used, including selection in rest-frame far-IR (e.g., Barger et al. 1998; Hughes et al. 1998; Blain et al. 1999; Eales et al. 2000), near-IR (e.g., Sawicki 2002), optical (e.g., Thompson et al. 1999; Cimatti et al. 2002; Sawicki et al. 2005), and the UV (e.g., Steidel et al. 1996, 1999, 2003, 2004; Sawicki et al. 1997; Lowenthal et al. 1997; Giavalisco 2002; Lehnert & Bremer 2003; Stanway et al. 2003; Iwata et al. 2003; Ouchi et al. 2004a). Among these different techniques, the Lyman break galaxy (LBG; Steidel et al. 1996, 2003) surveys have yielded the largest spectroscopically confirmed samples and the most active and detailed follow-up studies.

Follow-up observations of LBG samples have taught us much about the nature of these high-z galaxies, although, understandably, such follow-up has so far mainly focused on relatively luminous objects at $z \sim 3$, where the samples are largest and spectroscopy is easiest. We now know, for example, that LBGs are dominated by fairly young episodes of star formation and that they are enshrouded by large amounts of interstellar dust (e.g., Sawicki & Yee 1998; Ouchi et al. 1999; Shapley et al. 2001; Papovich et al. 2001; Vijh et al. 2003); that they are associated with massive dark matter halos (Adelberger et al. 1998; Giavalisco et al. 1998); that they have strong, starburst-driven outflows of material into the surrounding intergalactic medium (IGM; Pettini et al. 1998, 2001, 2002; Adelberger et al. 2003); and that they likely have subsolar, but not primordial, metallicities (Pettini et al. 2001, 2002). However, most LBG studies have focused primarily on relatively luminous objects ($L \gtrsim L^*$) at a single epoch ($z \sim 3$), and comparisons of LBG properties as a function of time and luminosity are still in their infancy.

Studying the properties of galaxies as a function of redshift has a straightforward motivation rooted in the fact that observing galaxies at different epochs allows us to study directly their evolution as a function of time. The initial search for and study of high-z galaxies has been motivated by the desire to find the progenitors of present-day galaxies, and recent comparisons of galaxy populations between different epochs beyond $z > 1$ are its direct and natural extension (e.g., Steidel et al. 1999; Adelberger et al. 2005; Ando et al. 2004; Ferguson et al. 2004; Papovich et al. 2004). Exploring the time domain holds obvious but important attractions.

The study of high-z galaxies as a function of their luminosity is less obvious to motivate, especially given the fact that high-z

[1] Based on data 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 NASA and was made possible by the generous financial support of the W. M. Keck Foundation.
[2] Current address: Department of Physics, University of California, Broida Hall, Building 572, Santa Barbara, CA 93106.

studies are observationally expensive even for luminous galaxies and far more so for galaxies that are intrinsically faint. Nevertheless, a wealth of potential information lies to be discovered in comparisons of galaxies as a function of luminosity. The galaxy luminosity function (LF) is not a simple power law as could be expected from the mass function of dark matter halos (e.g., Jenkins et al. 2001) but instead reflects the imprint of real differences in galaxy formation and evolution processes. If the luminosities of high-z galaxies correlate with the masses of their host dark matter halos, as is suggested by some clustering studies (Giavalisco & Dickinson 2001; Ouchi et al. 2004b), then the shape of the LF at high redshift likely bears the direct imprint of star formation-driven feedback as a function of halo mass. In addition, luminosity may also reflect the effects of processes such as fluctuating star formation rates (SFRs), whether induced by galaxy-galaxy interactions or by other mechanisms, or differences in the properties of interstellar dust. In any case, it is unlikely that galaxies of different luminosity are just trivially scaled copies of each other, and so to fully understand the story of galaxy formation, we must study not just the brightest, observationally most accessible members of the population, but also their fainter cousins. Such studies are particularly attractive because differences in evolution between galaxies with different UV luminosities (or, by extension, SFRs) can be expected to point us to some of the most relevant mechanisms responsible for driving galaxy evolution. Finally, for any reasonable LF, most galaxies are sub-L^* galaxies and most of the luminosity in the universe is contained in galaxies below L^*. By extension, so is most of the star formation and metal production activity. The hitherto neglected sub-L^* high-redshift galaxies deserve our avid attention for all of these important reasons.

One of the most basic descriptors of a galaxy population is its LF. The shape of the galaxy LF bears the imprint of galaxy formation and evolution processes. The characteristic break in the LF seen at both low and high redshift suggests that galaxies below L^* are not simple scaled replicas of those above L^* but differ from them in more substantial ways. At present relatively little is known about the shape of the faint end of the LF of galaxies at high redshift and about the evolution of the high-z LF. What studies have been done are limited by small, often single fields, such as the Hubble Deep Field (HDF), that can be affected by both sample and cosmic variance (e.g., Sawicki et al. 1997; Steidel et al. 1999) or use samples whose fidelity has not been well tested with spectroscopy (e.g., Iwata et al. 2003; Ouchi et al. 2004a; Gabasch et al. 2004).

In this paper we use our large, very faint Keck Deep Field (KDF) galaxy samples to construct the LFs of high-z star-forming galaxies over a wide span of cosmic time (0.6, 0.8, and 0.8 Gyr from $z \sim 4$ to $\sim$3 to $\sim$2.2 to $\sim$1.7, respectively) and reaching to very faint limits ($\mathcal{R} = 27$, or $M_{1700} \sim -18$, equivalent to SFR of $\sim 1\ M_\odot\ \mathrm{yr}^{-1}$). At $z \sim 4$ and 3, we combine our samples with LF measurements by Steidel et al. (1999) made from shallower but larger area surveys to study the rest-frame UV-selected galaxy LF over up to a factor of 100 in luminosity.

This paper is structured as follows. In § 2 we briefly describe our KDF sample of faint, UV-selected galaxies. In § 3 we describe the details of how we calculate the LF, and in § 4 we describe our results and focus on examining the possible sources of systematic error that may affect them. In § 5 we examine several intriguing evolutionary trends in the LF. In § 6 we discuss some possible interpretations of the observed evolution and also point out the potential new approaches to the study of galaxy evolution at high redshift that the evolving LF opens to us. Finally, in § 7 we summarize our results. As in all of the papers in the KDF series, we use the AB flux normalization (Oke 1974) and adopt $\Omega_M = 0.3$, $\Omega_\Lambda = 0.7$, and $H_0 = 70\ \mathrm{km\ s^{-1}\ Mpc^{-1}}$.

2. THE DATA

Our study of the faint end of the high-redshift LF uses data from our very deep $U_nG\mathcal{R}I$ Keck imaging survey, the KDF. The KDF was specifically designed to explore the evolution of the population of very faint (sub-L^*) star-forming galaxies at high redshift, including their LF and luminosity-dependent clustering. The KDF probes the hitherto poorly explored faint end of the galaxy population at redshifts $z \sim 4-1.7$ by extending to fainter magnitudes the well-known color selection techniques used by Steidel et al. (1999, 2003, 2004). The KDF survey is described in detail in the companion paper by Sawicki & Thompson (2005, hereafter KDF I), which gives a detailed description of the observations, data reduction, and selection of star-forming galaxy samples at $z \sim 2-4$. Here we give only a brief overview of its main characteristics.

The KDF survey uses the very same $U_nG\mathcal{R}I$ filter set and color-color selection criteria that are used so successfully by Steidel et al. (1999, 2003, 2004) to select their high-redshift samples, but it reaches a limiting magnitude $\mathcal{R}_{\rm lim} = 27$, which is 1.5 mag deeper than the surveys by the Steidel group. Because of our use of an identical $U_nG\mathcal{R}I$ filter set, our KDF data are a direct and straightforward extension to fainter magnitudes of the spectroscopically tested and well-understood samples of Steidel et al. (1999, 2003, 2004). Because of the extensive spectroscopic work of the Steidel team, selection effects, including foreground interloper fractions (which are only a few percent), are well known and can be safely used for our fainter samples as explained in KDF I.

The KDFs cover a total area of 169 arcmin^2 and consist of five fields (called fields 02A, 03A, 03B, 09A, and 09B) that were observed separately and so reach slightly different depths ranging over $\mathcal{R}_{\rm lim} \sim 26.7-27.3$ (50% completeness). The five fields are grouped into three patches (patches 02, 03, and 09) that are spatially well separated on the sky. The division of the KDF into these three spatially independent patches gives us the extremely important ability to monitor the magnitude and impact of cosmic variance.

Our ground-based images have typical seeing of $\sim 1''$, ensuring that high-z galaxies are unresolved and can be treated as point sources. This relatively poor spatial resolution is a blessing in disguise as it drastically reduces concerns about galaxy size selection biases that are present in observations with better image quality, such as *Hubble Space Telescope* (*HST*) data.

Photometry is done in a manner virtually identical to that of the Steidel et al. (1999, 2003, 2004) work, namely, with object detection in very deep $\mathcal{R}$-band images and color measurement through matched $2''$ diameter apertures on images smoothed to a common seeing. Our $U_nG\mathcal{R}I$ filter set, the same as that used by the Steidel team, allows us to select high-z galaxies in a manner that is identical to their brighter samples (for details of the selection criteria see KDF I or Steidel et al. 1999, 2003, 2004). The photometric completeness of our survey, tested carefully using simulations, is similar at $\mathcal{R} = 27$ to that of the Steidel et al. (1999, 2003, 2004) surveys at $\mathcal{R} = 25.5$. The KDF therefore probes a factor of 4 deeper in luminosity than the work of the Steidel team.

To its nominal completeness limit of $\mathcal{R}_{\rm lim} = 27$, the KDFs contain 427 $G\mathcal{R}I$-selected $z \sim 4$ LBGs, 1481 $U_nG\mathcal{R}$-selected $z \sim 3$ LBGs, 2417 $U_nG\mathcal{R}$-selected $z \sim 2.2$ star-forming galaxies, and 2043 $U_nG\mathcal{R}$-selected $z \sim 1.7$ star-forming galaxies.

3. CALCULATION OF THE LUMINOSITY FUNCTION

We use the effective volume, $V_{\rm eff}$, approach to compute the LFs. Our approach is virtually identical to that used by Steidel

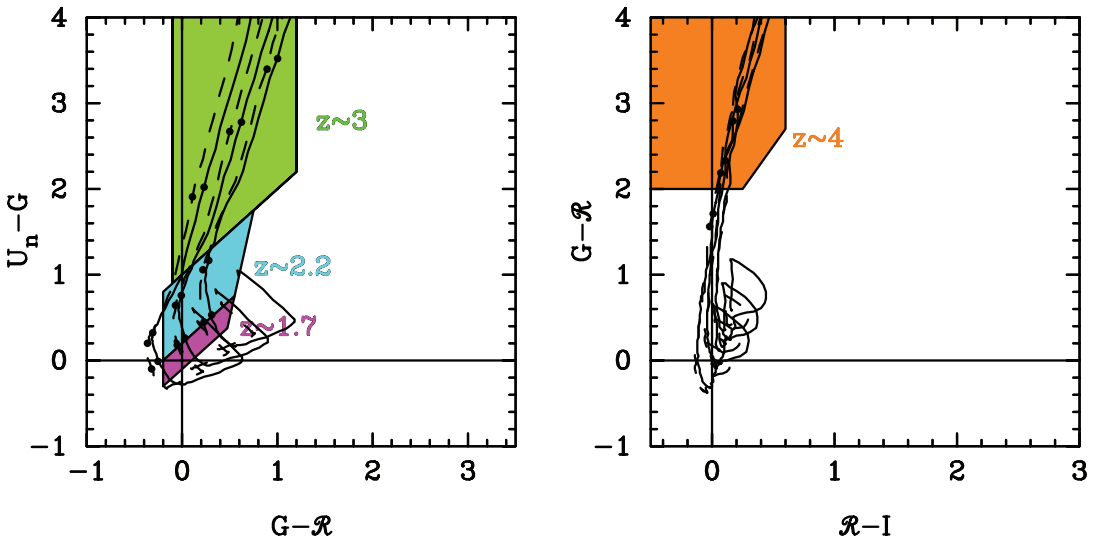

FIG. 1.—Colors of model galaxies in the $U_nG\mathcal{R}$ (*left*) and $G\mathcal{R}I$ (*right*). The filled regions represent the color-color selection criteria used to select our samples of high-z galaxies (see KDF I; Steidel et al. 1999, 2003, 2004). The tracks show model colors of 100 Myr old (*solid lines*) and 10 Myr old (*dashed lines*) starbursts for three values of reddening each, $E(B-V) = 0$, 0.15, and 0.3. Reddening generally increases from lower left to upper right. The points mark the location of $z = 1.7$, 2.2, 3, and 4 on each of the tracks.

et al. (1999) on their brighter $z \sim 3$ and $\sim$4 samples. This section of the paper is devoted to a detailed discussion of the technique we use to calculate the LF. We defer the discussion of our actual LF results to § 4 and later.

The $V_{\rm eff}$ approach to calculating the LF is straightforward. In brief, (1) for each redshift sample ($z \sim 4$, $\sim$3, $\sim$2.2, and $\sim$1.7) we first use simulations to determine the effective volumes, $V_{\rm eff}$, of the survey as a function of apparent magnitude, i.e., volumes that account for incompleteness due to objects missing from the sample. (2) We then combine these $V_{\rm eff}$ with the observed galaxy counts to compute the incompleteness-corrected number density of galaxies at each redshift as a function of apparent magnitude. (3) Finally, we convert these apparent magnitude number densities into absolute magnitude ones to arrive at the LFs. These three steps are described in detail in §§ 3.1, 3.2, and 3.3, respectively.

We measure the LF at rest-frame 1700 Å for two reasons: because (1) LF calculation at this rest-frame wavelength matches the LF analysis by Steidel et al. (1999) using brighter $z \sim 3$ and $\sim$4 LBG samples, but also because, as seen in more detail in § 3.3, (2) rest-frame 1700 Å very closely matches the observed-frame I band at $z \sim 4$, $\mathcal{R}$ band at $z \sim 3$, and G band at $z \sim 1.7$, thereby nearly eliminating uncertainties in k-corrections. At $z \sim 2.2$, 1700 Å is located between the G and $\mathcal{R}$ bands, and so for our $z \sim 2.2$ objects we construct composite $G\mathcal{R}$ magnitudes that more closely match rest-frame 1700 Å than do either G band or $\mathcal{R}$ band alone. Our composite $G\mathcal{R}$ magnitudes are a simple average of the G- and $\mathcal{R}$-band fluxes,

$$G\mathcal{R} = -2.5 \log\left(\frac{10^{-0.4G} + 10^{-0.4\mathcal{R}}}{2}\right), \qquad (1)$$

with uncertainties calculated by combining the G- and $\mathcal{R}$-band uncertainties in quadrature.

3.1. *Calculating the Effective Volumes of the Survey*

We must first compute the effective volume of the survey, as a function of apparent magnitude, for objects in each redshift sample. Because galaxies scatter in and out of the color-color selection regions (see Fig. 1 here and Figs. 4 and 5 in KDF I) that we use to select our high-z galaxy samples, and because this scattering will depend on the size of photometric uncertainties, we must compute our effective volumes as a function of apparent magnitude. Our calculation of effective volumes is accomplished through simulations in which we implant, and then seek to recover, artificial objects with colors and magnitudes representative of star-forming high-z galaxies. The description of these simulations is the subject of this section.

3.1.1. *Modeling the Colors of High-z Galaxies*

As the first step, we calculate a grid of model colors expected of high-z galaxies. We start with model spectral energy distributions (SEDs) of star-forming galaxies from the 1996 version of the Bruzual & Charlot (1993) spectral synthesis library. We use the continually star-forming models with solar metallicity and Salpeter initial mass function (IMF). We next redden them with a set of extinction values in the Calzetti (1997) starburst dust prescription. We then complete the redshift dimension of the grid by stretching these SEDs by $(1 + z)$ and attenuating them using the Madau (1995) prescription for continuum and line blanketing due to intergalactic hydrogen along the line of sight. Finally, we integrate the resultant reddened, observer-frame model spectra through the $U_nG\mathcal{R}I$ filter transmission curves to arrive at the predicted colors of high-z star-forming galaxies. Some examples of model galaxy colors are shown in Figure 1, where they are overplotted on top of regions of color-color space used to define our galaxy samples.

The largest influence on the colors of high-z galaxies is wielded by, first, attenuation by intergalactic hydrogen gas blueward of the Lyman break and, second, reddening due to interstellar dust internal to the galaxies. Other variables such as age, star formation history, stellar IMF, or metallicity can also play a role, but their effects are small in comparison, and especially so at the rest-frame UV wavelengths that concern us in this study. Consequently, we adopt fixed values for most of these parameters and explore only how our results vary with the adopted reddening and starburst age.

For practical reasons, we must restrict our choices to a limited set of these parameters. We are guided in our choice of dust

attenuation and starburst age by the observed values of these quantities in $z \sim 3$ LBGs. Early on, Sawicki & Yee (1998) studied the rest-frame UV through optical broadband photometry of 17 spectroscopically confirmed $z \sim 3$ LBGs in the HDF and concluded that these objects are dominated by young stellar populations ($\lesssim$0.2 Gyr) and substantial amounts of dust [median $E(B-V) \sim 0.3$]. However, the bulk (11/17) of the objects in their analysis came from the spectroscopic sample of Lowenthal et al. (1997), who allowed objects that are redder, and so presumably more dusty, than those selected using the now so familiar criteria of Steidel et al. (1999, 2003, 2004) that are used in our present KDF work. Indeed, Shapley et al. (2001) applied the SED fitting technique of Sawicki & Yee (1998) to a large sample of $z \sim 3$ LBGs selected solely using the Steidel et al. (1999, 2003, 2004) selection criteria and found a median $E(B-V) = 0.16$ (lower than Sawicki & Yee 1998) and concluded that LBGs undergo relatively short periods (50–100 Myr) of very intense star formation followed by a more quiescent star-forming phase (see also Sawicki & Yee 1998).

It is thus clear that at least bright ($\mathcal{R} \lesssim 25$) LBGs at $z \sim 3$ are dominated by fairly short episodes of star formation and significant dust obscuration. It remains unclear whether this is also the case at lower and higher redshifts and at the fainter magnitudes that we reach in the KDF. Nevertheless, motivated by the results of Sawicki & Yee (1998) and Shapley et al. (2001), we take as our fiducial model the 100 Myr old star-forming SED from the 1996 version of the Bruzual & Charlot (1993) spectral synthesis library and attenuate it with $E(B-V) = 0.15$ of dust. In § 3.1.2 we show that this fiducial model reproduces the Steidel et al. (1999, 2003, 2004) observed redshift distributions of $z \sim 2.2$, $\sim$3, and $\sim$4 galaxy samples (the case for $z \sim 1.7$ is less clear). Nevertheless, to monitor the impact of our choice of SED model on our LF results, we carry out all of our calculations in parallel, considering a grid of SED models that includes two stellar population ages, 10 and 100 Myr, and seven values of dust attenuation, $E(B-V) = 0-0.3$ in steps of 0.05. As we discuss in § 4.3, the dependence of the LF on these assumed dust and age values is negligibly small at $z \sim 4$ and $\sim$3 but becomes more significant at the lower redshifts.

3.1.2. *The Sample Completeness Function $p(m,\ z)$*

Next, we must correct for incompleteness of our catalogs that is brought on by both the imperfect object detection efficiency and the scatter of high-z galaxies across the boundaries of our color-color selection boxes. We do not (here or elsewhere in our LF calculation) explicitly correct for foreground, low-z interlopers that contaminate our high-z sample. Such interloper contamination is known to be very small in the spectroscopic samples of Steidel et al. (1999, 2003, 2004). Because, as is discussed in KDF I, the contamination fraction is expected at worst to remain constant and in all likelihood to fall toward fainter magnitudes, the contamination should be equally small or even smaller in our KDF data.

We measure the amount of incompleteness by implanting simulated galaxies into our images and then seeking to recover them using the very same procedures that we used in making our data catalogs (see KDF I). Incompleteness (both detection incompleteness and the loss of galaxies due to scattering out of the high-z color-color selection boxes) is a function of apparent magnitude, with fainter galaxies suffering larger incompleteness than their brighter kin. It also depends on the true colors of our target galaxies and hence their redshifts and intrinsic SEDs. We insert artificial objects with the expected colors of high-z galaxies generated as described in § 3.1.1. Our artificial objects have point-source profiles because, as we discussed in KDF I, the seeing in our images is sufficiently poor (FWHM $\sim 1''$) to ensure that high-z galaxies are spatially unresolved. The recovered fractions form the completeness function $p(m, z)$, which is the probability that a galaxy of a given apparent magnitude (in I at $z \sim 4$, $\mathcal{R}$ at $z \sim 3$, $G\mathcal{R}$ at $z \sim 2.2$, and G at $z \sim 1.7$), redshift z, dust attenuation, and model age matches our sample selection criteria.

The function $p(m, z)$ is measured separately for each of the four redshift samples ($z \sim 4$, $\sim$3, $\sim$2.2, and $\sim$1.7) and, given the small differences in the image properties of our five KDF fields, is recalculated for each KDF field. The function is sampled in steps of $\Delta m = 0.5$ in input apparent magnitude and $\Delta z = 0.1$ in redshift and for the eight combinations of age and reddening discussed above. At each step in this parameter grid several hundred simulated objects are implanted at quasi-random locations in the image. These positions are always the same for the different steps in the parameter grid but are otherwise unremarkable and sample the images fairly.

We can use our calculated $p(m, z)$ to test whether our assumptions about age and amount of dust are reasonable, i.e., whether we can reproduce the observed redshift distributions of high-z populations. The shaded histograms in Figure 2 show the redshift distributions of the spectroscopic samples of Steidel et al. (1999, 2003, 2004). These Steidel et al. (1999, 2003, 2004) spectroscopic samples contain galaxies with a range of apparent magnitudes and represent the underlying color-selected population convolved with a spectroscopic success function that is not trivial to model. However, the bulk of their spectroscopic samples consists of galaxies with $\mathcal{R} \sim 25$, i.e., galaxies with photometric errors that are similar to those of $\mathcal{R} \sim 26.5$ galaxies in the KDF.

Thus, our $p(m, z)$ for $\mathcal{R} = 26.5$ KDF galaxies should match the observed redshift distributions if our modeling is a reasonable representation of reality. Figure 2 shows that this is indeed the case for our fiducial model with $E(B-V) = 0.15$ and starburst age of 100 Myr (the 10 Myr models, which are not shown, are also very good): within each panel of Figure 2, the solid lines show the function $p(\mathcal{R} = 26.5,\ z)$ for the three values of $E(B-V)$ (which generally increase from left to right), with the thicker line marking $E(B-V) = 0.15$. Although our fiducial model does not give a uniquely matching solution [for example, a superposition of lower and higher $E(B-V)$ values might work just as well], it does give *a* good agreement with the data. The fiducial model works remarkably well at $z \sim 4$, 3, and 2.2; it works less well at $z \sim 1.7$ as it predicts a redshift distribution with a somewhat lower median redshift than is observed, but so do all of our other $p(m, z)$ at that redshift (we revisit this issue later). At any rate, the ability of our $p(m, z)$ modeling to reproduce the observed redshift distributions gives us confidence in that modeling, in the modeling of effective survey volumes, $V_{\rm eff}$, that are based on it in the next subsection, and thence in our estimate of the LFs.

3.1.3. *The Effective Volumes*

Finally, $V_{\rm eff}$ is calculated for each of the five fields by integrating the probability function $p(m, z)$ over redshift:

$$V_{\rm eff}(m) = A_f \int \frac{dV}{dz} p(m,\ z)\, dz, \tag{2}$$

where dV/dz is the comoving volume per square arcminute in redshift slice dz at redshift z and A_f is the area of the field in arcminutes. The $V_{\rm eff}$ is calculated separately for each color-selected redshift sample ($z \sim 4$, $\sim$3, $\sim$2.2, $\sim$1.7), each of the galaxy models (which consist of the eight combinations of reddening

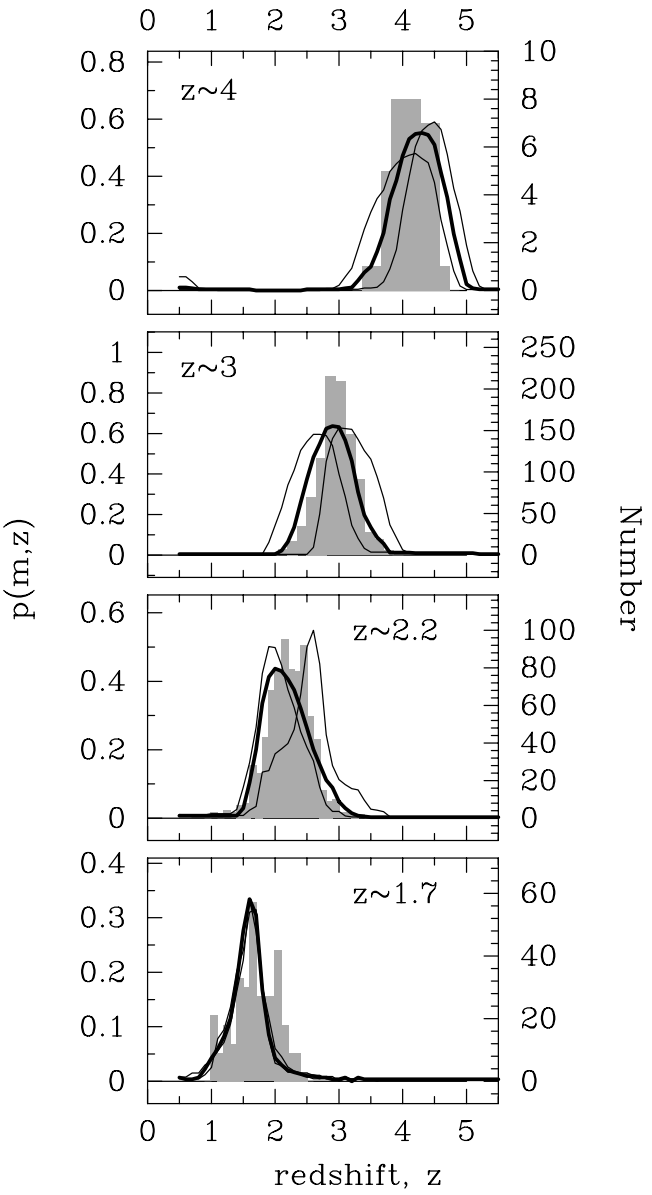

FIG. 2.—Comparison of our $p(m, z)$ models with the observed redshift distributions of UV-selected high-z galaxies. The shaded histograms show the spectroscopic redshift distributions of Steidel et al. (1999, 2003, 2004), while the solid lines show our $p(m, z)$ for $E(B-V) = 0$, 0.15, and 0.3. The $E(B-V) = 0.15$ fiducial model is marked with a thick line and, generally, in each panel the peak of the $p(m, z)$ moves to lower redshifts with increasing $E(B-V)$. These $p(m, z)$ models are for $\mathcal{R} = 26.5$, a magnitude at which typical photometric errors in the KDF correspond to those in the observed spectroscopic samples whose photometry is shallower than ours. Here results for our 09A field are shown, but the other four fields give similar lines that show good agreement between our fiducial model and the observed redshift distribution.

and starburst age), each apparent magnitude step ($\Delta m = 0.5$ in I, $\mathcal{R}$, $G\mathcal{R}$, or G, depending on the redshift sample being considered), and each of the five KDF fields.

3.2. *The Incompleteness-corrected Number Counts*

We next calculate the incompleteness-corrected galaxy density $\phi(m)$ as a function of apparent magnitude and for each $V_{\rm eff}$ model. We carry out the calculation separately for each of the redshift samples, $z \sim 4$, $\sim$3, $\sim$2.2, and $\sim$1.7. The calculation is first carried out independently for each KDF field and then the results are averaged together.

For each field f we compute in 0.5 mag bins the number of galaxies satisfying the color selection criteria and then correct for incompleteness using the effective volume:

$$\phi_f(m) = 2\frac{N_f(m)}{V_{\rm eff}(m)}. \quad (3)$$

Here $N_f(m)$ is the number of observed galaxies within the magnitude bin $m \pm 0.25$ in that field, $V_{\rm eff}(m)$ is the effective volume of that field, $\phi_f(m)$ are the incompleteness-corrected number counts in units of mag^{-1} Mpc^{-3}, and the factor of 2 converts from counts in the 0.5 mag bins to counts mag^{-1}.

The results of the individual fields are then weighted by field area, A_f, and averaged to yield the incompleteness-corrected galaxy number density for the entire KDF survey,

$$\phi(m) = \frac{\Sigma_f A_f \phi_f(m)}{\Sigma_f A_f}. \quad (4)$$

We restrict the calculation of the final KDF $\phi(m)$ to magnitude bins no fainter than the 50% detection completeness limit in each field of the KDF. This limit is deeper by 0.5 mag for three of our fields than for the other two (see § 2 and KDF I), and so the average $\phi(m)$ in equation (4) is computed using all five fields at all magnitudes except for the faintest magnitude bin, in which only the deeper three fields are used.

The uncertainty in $\phi(m)$ combines two sources of uncertainty, namely, the uncertainty in number statistics, $\delta\phi_N(m)$, and an estimate of field-to-field fluctuations, $\delta\phi_{f2f}(m)$. The number statistics uncertainty is simply the Gaussian $\delta\phi_N(m) = \phi(m)[N(m)]^{-0.5}$, where $N(m)$ is the total number of galaxies in all five (or three) fields in that magnitude bin. The field-to-field uncertainty $\delta\phi_{f2f}$ is estimated using bootstrap resampling, whereby we generate 500 new realizations of $\phi(m)$ via equation (4), but now in each realization choosing fields randomly with replacement and so allowing the same field to be included more than once (or not at all) in a given realization. The field-to-field uncertainty $\delta\phi_{f2f}$ is then taken to be the rms value of the 500 $\phi(m)$-resampled realizations. These two sources of uncertainty are then added in quadrature $\delta\phi = [(\delta\phi_N)^2 + (\delta\phi_{f2f})^2]^{0.5}$ to give us the total uncertainty.

3.3. *Absolute Magnitudes*

The final step is to convert the $\phi(m)$ into the LF $\phi(M)$. As we discussed earlier, we compute the LF at rest-frame 1700 Å both to minimize k-corrections and to retain commonality with the work of Steidel et al. (1999) at brighter magnitudes.

The absolute magnitude, M, is derived using the usual cosmological distance modulus, DM, and k-correction, K,

$$M_{1700} = m_{\lambda_{\rm obs}} - {\rm DM} - K, \quad (5)$$

which we rewrite as

$$M_{1700} = m_{\lambda_{\rm obs}} - 5\log(D_L/10\ {\rm pc}) + 2.5\log(1+z) + \left(m_{1700} - m_{\lambda_{\rm obs}/(1+z)}\right). \quad (6)$$

Here D_L is the luminosity distance and $m_{\lambda_{\rm obs}}$ is the observed magnitude in the principal filter for the redshift sample being considered (I, $\mathcal{R}$, $G\mathcal{R}$, and G for $z \sim 4$, 3, 2.2, and 1.7, respectively). The last term of equation (6), $(m_{1700} - m_{\lambda_{\rm obs}/(1+z)})$, is the k-correction color between rest-frame 1700 Å and the principal filter in the rest frame for the redshift sample in question. This k-correction color is expected to be very small because of our decision to work at rest-frame 1700 Å. This expectation is illustrated in the top panels of Figure 3, where we plot the k-correction color for

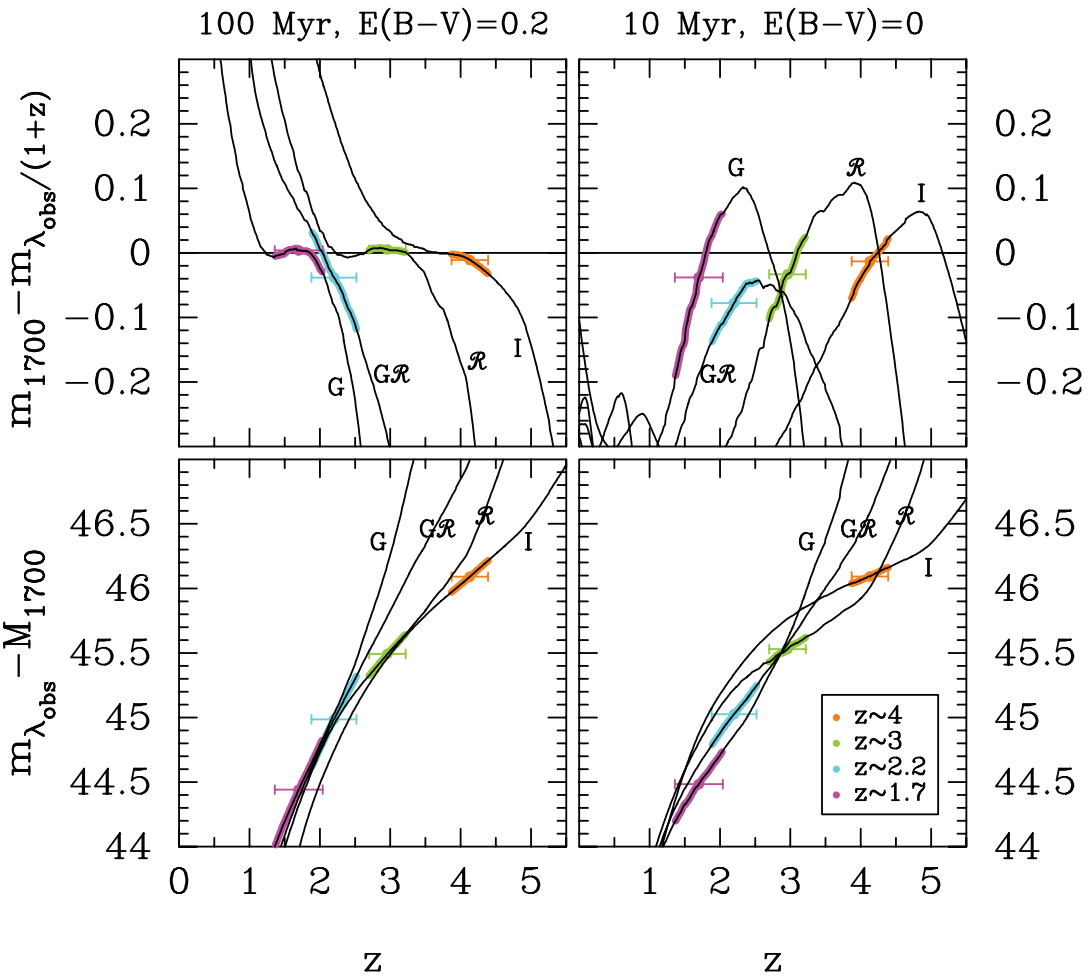

FIG. 3.—The *k*-correction color term (*top*) and the distance moduli (*bottom*) for two representative galaxy SEDs. Locations of the $z \sim 4$, ~3, ~2.2, and ~1.7 sample bin centers are marked. The top panels show the color terms that are needed to transform observed G, $G\mathcal{R}$, $\mathcal{R}$, and I magnitudes to rest-frame 1700 Å. The top left panel shows the color term for a 100 Myr old starburst with moderate dust attenuation, while the top right panel is for an unobscured 10 Myr starburst. The error bars show the FWHM redshift ranges ($\delta z \sim 0.3$; see KDF I for details) spanned by the color-color–selected samples, and the thick colored bands highlight the color term values corresponding to these redshift ranges. These plots illustrate that for the right choice of observed bandpass, namely, I for $z \sim 4$ LBGs, $\mathcal{R}$ for $z \sim 3$, $G\mathcal{R}$ for $z \sim 2.2$, and G for $z \sim 1.7$, the *k*-correction color is ~0. The bottom panels show the corresponding offsets between observed and absolute magnitudes (as defined in eq. [6], i.e., including the *k*-correction color term) and illustrate that the redshift uncertainty for our photometrically selected objects translates into only a small DM uncertainty and hence into only a small uncertainty in the derived absolute UV magnitude of the object, M_{1700}.

two representative galaxy models selected out of a larger ensemble that we tested. As Figure 3 illustrates, the value of the color term is indeed very close to zero for I band at $z \sim 4$, $\mathcal{R}$ band at $z \sim 3$, the composite $G\mathcal{R}$ at $z \sim 2.2$, and G band at $z \sim 1.7$. Over the redshift ranges selected by the color selection criteria we are using, the deviations from zero are typically no larger than 0.1 mag with a redshift-dependent δmag range of no more than roughly ± 0.1 for a given model. These small offsets are negligible, and so we set $(m_{1700} - m_{\lambda_{obs}/(1+z)})$ to zero in equation (6) to arrive at

$$M_{1700} = m_{\lambda_{obs}} - 5\log(D_L/10\ \mathrm{pc}) + 2.5\log(1+z). \quad (7)$$

Applying equation (7) to our $\phi(m)$, we at last arrive at the rest-frame 1700 Å LF, $\phi(M)$. The resulting LF measurements are described in § 4.

4. THE OBSERVED LUMINOSITY FUNCTIONS

4.1. *Description of the Luminosity Functions*

The data points in Figure 4 show our LF measured using our baseline $V_{\rm eff}$ model of a 100 Myr old starburst with $E(B-V) = 0.15$ (dependence on model assumptions is discussed in § 4.3). In addition to the KDF data, we also include the $z \sim 4$ and ~3 LF points of Steidel et al. (1999). We stress that our analysis of the KDF data follows closely the selection and LF analysis procedures used by Steidel et al. (1999) for these brighter galaxies, and so combining their results with ours should be robust and free of systematic uncertainties. The KDF and the Steidel et al. (1999) samples complement each other: the Steidel et al. (1999) measurements provide good statistics at the bright end but do not probe fainter than $M_{1700} = -20.5$ at $z \sim 3$ and $M_{1700} = -21$ at $z \sim 4$, while the KDFs have good statistics at the faint end, probing ~1.5 mag deeper, but lack the statistics at the bright end. Unless otherwise stated, throughout the rest of this paper we always combine the Steidel et al. (1999) measurements and our fainter KDF data when discussing the $z \sim 3$ and ~4 LFs.

To give a more physical meaning to our luminosity scale, we relate rest-frame UV magnitudes to SFRs. Combining the relation between SFR and luminosity given by Kennicutt (1998; see also Madau et al. 1998) with the definition of AB magnitudes (Oke 1974) gives

$$\mathrm{SFR} = 6.1 \times 10^{-(8+0.4M_{1700})}\ M_\odot\ \mathrm{yr}^{-1}. \quad (8)$$

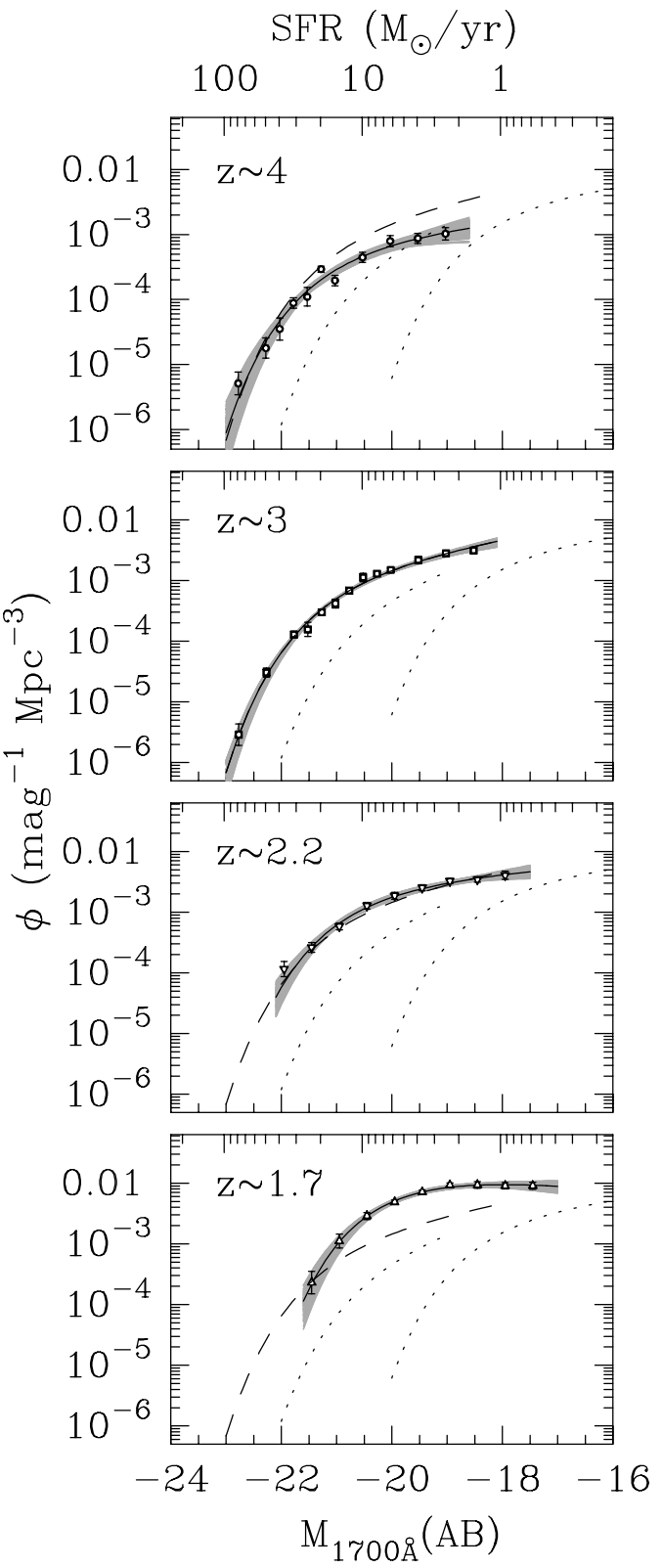

FIG. 4.—Rest-frame UV LFs at different redshifts. The solid lines show the best-fitting Schechter functions, while the shaded bands show the corresponding 68.3% confidence region. As is described in the text, the error bars include both the $N^{1/2}$ statistics and a bootstrap estimate of field-to-field variations. The dashed fiducial line simply reproduces the $z \sim 3$ line. The dotted lines show the *GALEX* rest-frame 1500 Å LFs for comparison; the rightmost, fainter one is for $z \sim 0$ and the leftmost, brighter one is for $z \sim 1$. As we discuss in the text, we regard the $z \sim 4$ and 3 LFs shown here to be highly trustworthy, the $z \sim 2.2$ to represent a firm lower limit on the galaxy number density, but the $z \sim 1.7$ to be questionable due to systematic biases.

This conversion is valid in the absence of dust and for a stellar population that is forming stars continuously and whose UV light is dominated by massive, short-lived stars that are being produced on a Salpeter (1955) stellar IMF with mass range $0.1 \leq M_\odot \leq 100$. The SFR scales are plotted on the top axes in Figure 4. The SFR scale should not be taken too literally because of the uncertainties in the assumptions underlying equation (8). Nevertheless, this scale gives us a useful, more physical reference frame for the LFs. As Figure 4 shows, the KDF LF reaches to $M_{1700} = -18$ at $z \sim 2.2$ and $M_{1700} = -19$ at $z \sim 4$. It thus rivals in depth LFs measured in the HDFs (e.g., Sawicki et al. 1997; Steidel et al. 1999). Under the aforementioned assumptions, our KDF LF reaches down to galaxies with SFRs of $\sim 2\ M_\odot\ \mathrm{yr}^{-1}$ and lower, comparable to the SFR in the Milky Way today.

The solid lines in each panel of Figure 4 show Schechter function fits to the data (Schechter 1976), and the gray shaded regions show the 1 σ uncertainties in those fits. We defer the description of the details of the Schechter function fitting to § 4.2, and here we use the fits only to guide the eye and to give a first comparison between LFs at different redshifts.

Our $z \sim 3$ LF is the best constrained of all of the redshift bins we consider, and so, when comparing LFs at different redshifts, we use the $z \sim 3$ LF as reference. The dashed lines in the other panels of Figure 4 show the fit to this $z \sim 3$ fiducial. Comparing the data at $z \sim 4$ with our fiducial $z \sim 3$ lines immediately suggests that the LF undergoes evolution with redshift: the number density of faint galaxies at $z \sim 4$ appears to be significantly lower than at $z \sim 3$. At the same time, the number density of luminous galaxies appears to remain unchanged from $z \sim 4$ to ~ 3. In §§ 4.3 and 5 we explore in detail whether this LF evolution is real or simply the result of a selection effect or other artifacts (we conclude that it is very likely real at least between $z \sim 4$ and ~ 3).

For comparison, Figure 4 also shows the recent low-z rest-frame 1500 Å LF measurements from the *GALEX* mission. These LFs are shown as dotted lines in Figure 4, with the left dotted line showing the $z \sim 1$ *GALEX* LF and the right one showing the $z \sim 0$ LF (Arnouts et al. 2005 and Wyder et al. 2005, respectively). The strong evolution of the LF of star-forming galaxies from $z \sim 0$ to ~ 1 seen in the *GALEX* data has been recognized for some time (e.g., Lilly et al. 1995) and is responsible for the steep rise in the UV comoving luminosity density of the universe over that redshift interval (e.g., Lilly et al. 1996; Schiminovich et al. 2005). Similarly, the increase in the number of luminous galaxies from $z \sim 1$ to higher redshifts, $z \gtrsim 2$, is part of the familiar UV sketch of the cosmic star formation history (e.g., Madau et al. 1996; Sawicki et al. 1997; Giavalisco et al. 2004). In contrast to these well-known broad trends, the more subtle evolution in the LF from $z \sim 4$ through ~ 3 to ~ 2.2 that is revealed in our KDF data has not been so far explored because until now there was a lack of well-selected and well-tested samples that have good statistics over a wide range in luminosity.

4.2. *Parametric Representation of the LF*

We fit the binned LF data with the Schechter (1976) function,

$$\begin{aligned}\phi_{\mathrm{model}}(M) &= \phi^* \hat{\phi}(M) \\ &= \phi^* 0.4 \ln(10) \mathrm{dex} \left\{ [0.4(M^* - M)]^{(1+\alpha)} \right\} \\ &\quad \times \exp\left(-10^{0.4(M^*-M)}\right), \end{aligned} \tag{9}$$

TABLE 1
PARAMETERS OF SCHECHTER FUNCTION FITS

z	Steidel Type[a]	M^*_{1700}	$\log \phi^{*}$[b]	α
1.7.....	BM	$-19.80^{+0.32}_{-0.26}$	$-1.77^{+0.11}_{-0.11}$	$-0.81^{+0.21}_{-0.15}$
2.2.....	BX	$-20.60^{+0.38}_{-0.44}$	$-2.52^{+0.20}_{-0.26}$	$-1.20^{+0.24}_{-0.22}$
3.....	C, D, M, and MD	$-20.90^{+0.22}_{-0.14}$	$-2.77^{+0.13}_{-0.09}$	$-1.43^{+0.17}_{-0.09}$
4.....		$-21.00^{+0.40}_{-0.46}$	$-3.07^{+0.21}_{-0.33}$	$-1.26^{+0.40}_{-0.36}$

[a] In the nomenclature of Steidel et al. (2003, 2004).
[b] In units of Mpc^{-3}.

which we evaluate over M^*, ϕ^*, and α. In practice, we do the fitting using χ^2 minimization, where for a grid of M^* and α values we compute a corresponding grid of χ^2 values using

$$\chi^2(M^*,\ \phi^*,\ \alpha) = \sum_M \left[\frac{\phi_{\rm data}(M) - \phi_{\rm model}(M)}{\sigma(M)}\right]^2, \quad (10)$$

where the sum is taken over the M_{1700} magnitude bins and $\phi_{\rm model}(M)$ are computed using equation (9). Instead of adding a third (ϕ^*) dimension to the grid, the computation is considerably accelerated by optimally calculating ϕ^* at each (M^*, α) in the grid using the analytic relation

$$\phi^* = \frac{\sum_M \hat{\phi}(M)\phi_{\rm data}(M)/\sigma^2(M)}{\sum_M \hat{\phi}^2(M)/\sigma^2(M)}, \quad (11)$$

which is derived by minimizing equation (10) via $\partial\chi^2/\partial\phi^* = 0$. We then search the grid of χ^2 values to select its minimum, $\chi^2_{\rm min}$, and adopt its corresponding M^*, ϕ^*, and α as the best-fit Schechter function parameters. The values of these best-fit parameters are listed in Table 1 and plotted in Figure 5.

The error contours shown in Figure 5 are computed by recalculating the best-fitting M^*, ϕ^*, and α, but now with sets of data values $\phi_{\rm data}(M)$ that have been perturbed randomly according to their standard deviations $\sigma(M)$. For each of the redshift bins, we generate 250 such perturbed realizations and use their χ^2 to map out the regions of parameter space that correspond to the best-fitting 68.3% of such realizations.

4.3. *Systematic Effects in the LF Measurement*

4.3.1. *Uncertainty due to k-Corrections*

As we discussed in § 3.3, setting the *k*-correction color term to zero introduces only a very small, $\lesssim$0.1 mag, systematic bias in the determination of absolute magnitude (see Fig. 3). Consequently, $\sim$0.1 mag can be taken as the systematic uncertainty in our determination of the positions of $\phi(M)$ bins. This uncertainty is too small to affect our LFs significantly, and so we do not consider it further.

The lack of spectroscopic redshift information for objects in our sample may also introduce a systematic effect, albeit, as we show here, a negligible one. Our lack of spectroscopic redshifts means that we do not know whether a particular object in a given color-selected redshift sample is near the lower or the higher end of the redshift interval. This uncertainty may introduce a systematic bias since at the same apparent magnitude intrinsically less luminous objects are more likely to come from somewhat lower redshifts (and hence occupy a smaller effective volume) than intrinsically more luminous galaxies. However, as is shown in the bottom panels of Figure 3, the effect of this redshift uncertainty on the value of the distance modulus DM and hence on the derived M_{1700} is small: it is $\pm$0.2 mag for the $z \sim 4$ and $\sim$3 samples,

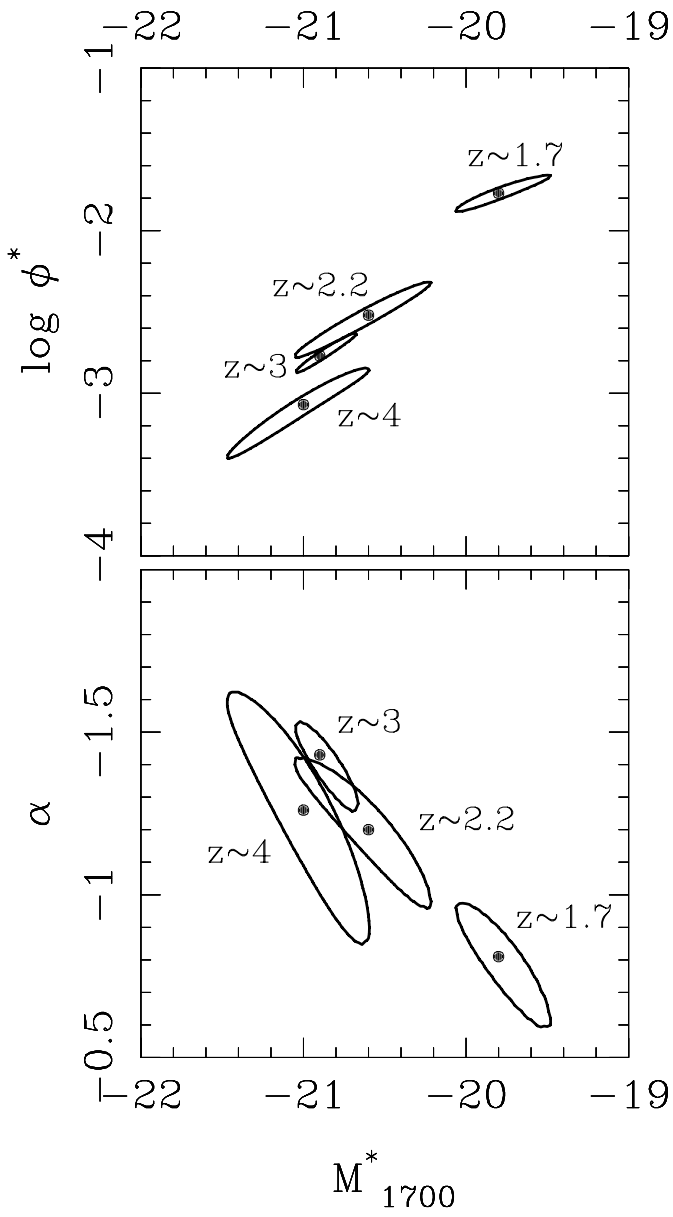

FIG. 5.—Schechter LF parameters and their 1 σ confidence regions.

$\pm$0.4 mag at lower redshifts. Such systematic offsets of a few tenths of a magnitude in M_{1700} are comparable to uncertainties introduced into the LF measurement by $N^{1/2}$ statistics and field-to-field variance as reflected, for example, in the uncertainties shown in Figure 5. Moreover, because the shapes of the LFs are similar in all redshift bins, the systematic bias introduced by this effect works in the same direction at $z \sim 4$, 3, 2.2, and 1.7, thereby reducing any systematic differences between the redshift bins. In summary, then, the biases introduced by our lack of spectroscopic redshifts are highly unlikely to drastically affect our results.

4.3.2. *Dependence on Cosmology*

Throughout this paper we have assumed a cosmological model with $(\Omega_M, \Omega_\Lambda, H_0) = (0.3, 0.7, 70\ \mathrm{km\ s^{-1}\ Mpc^{-1}})$. However, it is well known that many quantities, including the luminosities and volume elements directly relevant in LF calculation, can depend strongly on the assumed cosmological model. Consequently, we feel it important to explore how the choice of cosmological model impacts our results.

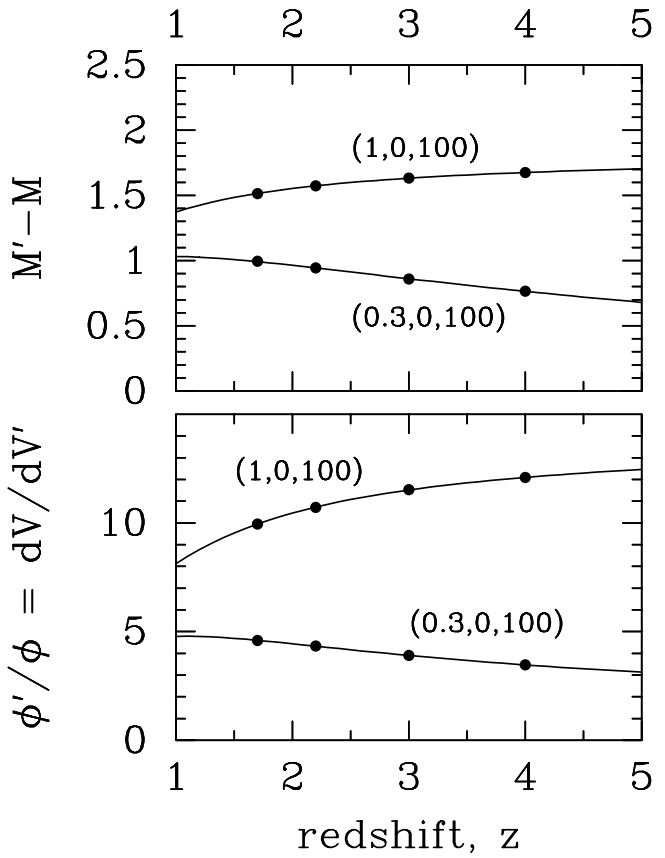

FIG. 6.—Effect of changing cosmology on absolute magnitudes (*top*) and number densities (*bottom*) as a function of redshift. The quantities plotted show the change incurred in transforming from the cosmology adopted in this paper, $(\Omega_M, \Omega_\Lambda, H_0) = (0.3, 0.7, 70\ \mathrm{km\ s^{-1}\ Mpc^{-1}})$, to the two alternative cosmologies labeled in the plots. The values of M and ϕ in the alternate cosmologies are denoted with primed quantities, while in our cosmology they are unprimed. While the change in M or ϕ at any given redshift can be quite large, the relative change from redshift to redshift is small, ensuring that, for reasonable cosmologies such as those considered here, any evolutionary trends seen in the LF are not subject to the assumed cosmology.

The choice of cosmology affects the derived LF primarily through a change in the relative number density normalization (i.e., a change in ϕ; eq. [3]) and luminosity (i.e., a change in the absolute magnitude scale, M; eq. [7]). The dependence of the faint-end slope is weak as it enters only through a small difference in how the change in cosmology affects the derived $V_{\rm eff}$ at different apparent magnitudes. We have tested that the effect on the faint-end slope is negligible by recomputing our LFs for a model with $(\Omega_M, \Omega_\Lambda, H_0) = (1, 0, 100\ \mathrm{km\,s^{-1}\ Mpc^{-1}})$, and consequently, we focus our discussion on the effects on M and ϕ.

Figure 6 shows how changing cosmology from our assumed $(\Omega_M, \Omega_\Lambda, H_0) = (0.3, 0.7, 70\ \mathrm{km\ s^{-1}\ Mpc^{-1}})$ to two other often-used cosmological models changes the resultant absolute magnitudes (*top panel*) and number densities (*bottom panel*). In both cases, the effect can be quite large: the difference in absolute magnitudes between the two models can be as large as 1–2 mag, and the change in number density as large as an order of magnitude. However, the relative changes in M and ϕ are only weakly dependent on redshift for a given cosmological model over the redshift range we study: relative to $z \sim 3$, the $z \sim 1.7$, $\sim$2.2, and $\sim$4 M scales are shifted by less than $\sim$0.15 mag by a change of our assumed cosmology; similarly, the number densities are also affected only weakly, at the $\lesssim$20% level. A direct consequence of this is that while the absolute ϕ and M scales depend strongly on cosmology, the relative shapes and ϕ and M normalizations from redshift bin to redshift bin are virtually unchanged. As a result, any real evolutionary trends seen in the LF from $z \sim 4$ to $\sim$3 to $\sim$2.2 to $\sim$1.7 are virtually independent of the assumed cosmology.

4.3.3. *Dependence on Galaxy Model Properties*

Our LF calculation relies, as it must, on our estimates of the effective volumes of the survey (see § 3). These $V_{\rm eff}$ estimates are based on models of the expected colors of high-z galaxies and therefore may depend critically on our model assumptions. Throughout this paper we have assumed a baseline model for high-z galaxies based on a 100 Myr old starburst reddened by $E(B-V) = 0.15$ of dust. As we have shown in § 3.1.2, this model reproduces well the observed redshift distributions of high-z galaxies, while models that are similar but have different amounts of dust do a poorer job. Nevertheless, we wish to explore how strongly model dependent are our $V_{\rm eff}$ estimates and the resultant LF measurements.

To test how model dependent our LF results are, we have repeated all of the steps in our LF calculation (§§ 3.1–4.2) for a grid of galaxy models that contains starbursts of two different ages (10 and 100 Myr) and eight values of $E(B-V)$ [0–0.3 in steps of 0.05, and also a composite model that contains a mixture of galaxies with $E(B-V)$ values drawn evenly from that $E(B-V)$ range]. We note that we have applied the different models to our own KDF points but were unable to do so to the points that come from Steidel et al. (1999) since we do not know in detail the $V_{\rm eff}$ model that was used in their LF work. Consequently, the effect of changing $V_{\rm eff}$ models will be seen only in our data and so will manifest itself fully at the faint end of the LFs, while the bright ends of the $z \sim 4$ and $\sim$3 LFs, which are dominated by the Steidel et al. (1999) data, will remain unchanged.

We show the effect of changing the $V_{\rm eff}$ model in two ways. First, the recomputed LFs are plotted in Figure 7, in which the data points at each M_{1700} show the results of seven LFs computed using the seven discrete $E(B-V)$ values, and the solid black lines show all of the corresponding Schechter function fits; the gray lines in the $z \sim 4$, $\sim$2.2, and $\sim$1.7 panels reproduce the $z \sim 3$ LF Schechter function fits for comparison. Second, the number density values of the Schechter function fits at a fixed magnitude, $M_{1700} = -20$, are shown as a function of $E(B-V)$ in Figure 8, where the black lines connect the 100 Myr old starburst models and the gray lines are for the 10 Myr old ones.

As can be seen in the $z \sim 3$ panel of Figure 7, the choice of $V_{\rm eff}$ galaxy model has a miniscule effect at $z \sim 3$: an effect of changing the $V_{\rm eff}$ model has an effect that is no larger than the errors due to number statistics and cosmic variance (see Fig. 4). Figure 8 confirms that at $z \sim 3$ the dependence on model is small: the number density of galaxies at $M_{1700} = -20$ ranges over the range of models by $\lesssim$5% compared to our baseline model. We therefore conclude that at $z \sim 3$ the LF we measure is very robust with respect to model assumptions.

While the results at $z \sim 3$ are virtually model independent, the choice of $V_{\rm eff}$ model has a larger (although still small) effect at $z \sim 4$. Here the $\phi_{M=-20}$ values range by up to roughly $\pm$20% with respect to our baseline model. However, despite this variation, the faint end of the $z \sim 4$ LF always remains substantially below the $z \sim 3$ LF, and the variation between different models at $z \sim 4$ is not larger than the uncertainties introduced by field-to-field variations and $N^{1/2}$ statistics, as can be seen by comparing Figures 7 and 4. We thus conclude that while the choice of $V_{\rm eff}$ model may possibly affect the details of the $z \sim 4$ LF results, it is unlikely to alter the qualitative trends seen between the $z \sim 4$ and $\sim$3 LFs.

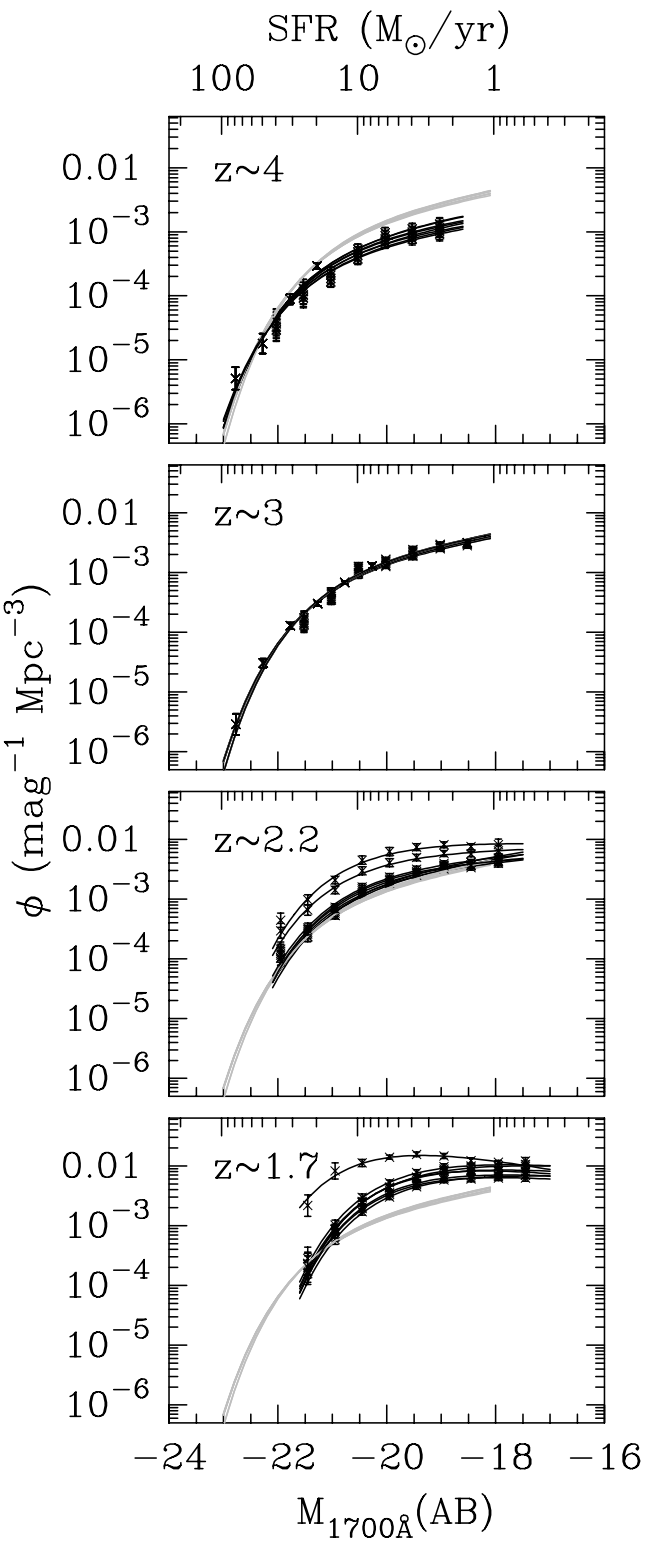

FIG. 7.—Dependence of the LF on the $V_{\rm eff}$ model. Variation in the derived LF for the range of considered $V_{\rm eff}$ models. The symbols and black lines show the results of recomputing the LFs using the seven different models of $V_{\rm eff}$ that result from seven different $E(B-V)$ values. The gray lines are the $z\sim3$ results replotted in the other redshift panels. See text for more details.

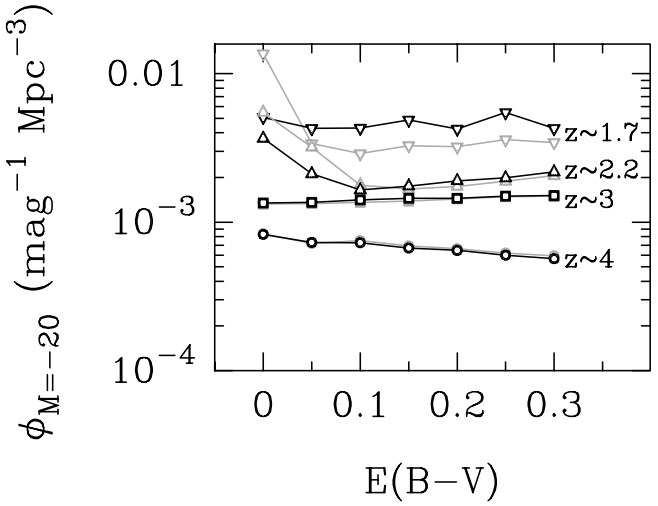

FIG. 8.—Dependence of the galaxy number density at $M_{1700}=-20$ on the $V_{\rm eff}$ model. The black symbols show 100 Myr old starburst models, and the gray ones are for 10 Myr old models.

The choice of $V_{\rm eff}$ potentially has the largest effect at $z\sim2.2$. While models with $E(B-V)>0$ give a tight grouping of LF results, the two models with $E(B-V)=0$ give LFs that have a substantially higher ϕ normalization. These higher LFs are a direct result of the fact that the colors of $z\sim2.2$ models with $E(B-V)=0$ approach very closely to the boundary of the color selection box (see Figs. 4 and 5 in KDF I): large numbers of model galaxies in these models are scattered out of the color selection box, resulting in low $p(m,z)$ values, low $V_{\rm eff}$, and hence high number densities ϕ in the resultant LFs. In our case the effect is not critical because (as we have argued in § 3.1.2) the $E(B-V)=0$ models are probably not realistic on other grounds. Moreover, even the adoption of the extreme $E(B-V)=0$ results does negate the possibility that the evolution seen from $z\sim4$ to ~3 continues onward to $z\sim2.2$; in fact, an $E(B-V)=0$ makes such an evolutionary trend stronger at $z\sim2.2$ than in the case of other $E(B-V)$ values. However, our $E(B-V)=0$ case at $z\sim2.2$ illustrates that in general one must be careful when computing LFs of color-selected samples because such LFs can be very strongly dependent on the assumed properties of the high-z galaxy populations that one is trying to study.

At $z\sim1.7$ the LF appears to have an even higher number density normalization than at higher redshifts. We note, however, that we do not place much faith in our determination of the $z\sim1.7$ LF, as we discuss in more detail in § 5.4.

In summary, we conclude that while the choice of $V_{\rm eff}$ models used in calculating the LF may have an effect on the LF results, in our case such effects are small: they are no larger than the uncertainties introduced by field-to-field variance and by $N^{1/2}$ statistics and in any case do not affect the qualitative evolutionary trends seen between $z\sim4$ and ~3. Meanwhile, the $z\sim2.2$ model $V_{\rm eff}$ may suffer from systematics that may adversely affect the accuracy of our LF measurements at those redshifts.

4.3.4. *Field-to-Field Fluctuations*

Galaxy clustering introduces field-to-field fluctuations in the galaxy distribution. This effect, often termed cosmic variance,[3] limits the accuracy with which the LF can be measured. This is particularly true for surveys that rely on small single pointings, such as the HDF (e.g., Sawicki et al. 1997; Steidel et al. 1999),

[3] This name is not strictly correct as cosmic variance refers to the variance in samples that are fundamentally limited by the finite size of the universe, such as, e.g., the largest scale fluctuations on the cosmic microwave background. Nevertheless, because it has entered the common nomenclature of galaxy evolution studies, we use the term cosmic variance interchangeably with field-to-field variance throughout this paper.

TABLE 2
NUMBER DENSITY $\log \phi^*$[a] IN DIFFERENT FIELDS OF THE SURVEY

z	Combined Fields[b]	02A	03A	03B	09A	09B
1.7.....	$-1.77^{+0.11}_{-0.11}$	−1.92	−1.80	−1.80	−1.70	−1.74
2.2.....	$-2.52^{+0.20}_{-0.26}$	−2.59	−2.41	−2.44	−2.65	−2.55
3.....	$-2.77^{+0.13}_{-0.09}$	−2.79	−2.69	−2.88	−2.77	−2.74
4.....	$-3.07^{+0.21}_{-0.33}$	−2.97	−3.24	−3.27	−3.16	−2.94

[a] In units of Mpc^{-3}.
[b] From fits to the full data as reported in Table 1.

the FORS Deep Field (FDF; Gabasch et al. 2004), or the Hubble Ultra Deep Field (e.g., Bouwens et al. 2006).

To mitigate the effects of cosmic variance, the KDF consists of five large fields grouped in three spatially independent patches. When calculating the LF, we have taken field-to-field variance into account by including a bootstrap resampling measurement of this variance in our error bars (see § 3.2). Here we illustrate the strength of field-to-field fluctuations by computing the characteristic galaxy number density ϕ^* in each of the five KDF fields. We make the measurement of the ϕ^* while holding the shape of the LF fixed. Specifically, at each redshift we hold α and M^* constant at the values we measured earlier for the full data set (Table 1) while letting ϕ^* be a free parameter. We also exclude the Steidel et al. (1999) $z \sim 4$ and $\sim$3 bright-end data from the measurement here. Table 2 summarizes the results and contrasts them with the ϕ^* values for the full data set. The field-to-field fluctuations are generally small. They are largest at $z \sim 4$ where the total number of objects in our sample is smallest: here the largest excursion, that in the 03B field, is $\sim$1.5 times the fiducial value, although the rms scatter is significantly smaller than that and consistent with our bootstrap estimates.

These field-to-field fluctuations, while relatively small in the KDF, underscore the need for multiple sight lines when determining the galaxy number density, LF, and derived quantities such as luminosity and SFR densities of the universe. Single-field studies, especially if they rely on *small* fields such as the HDF or the UDF, have no way of monitoring this important source of error. In contrast, here, in the multifield KDF, we have estimated field-to-field variance through bootstrap resampling and included it explicitly in our error budget.

4.4. *Comparison with Other Surveys*

Over the last few years several authors have presented LFs of UV-selected galaxies at high redshift, $z > 1$. The vast majority of these measurements were made using either full-blown photometric redshifts or the two-color selection techniques inspired by the success of the Steidel et al. (1999) surveys. The bulk of this work can be divided into two groups: those that use the very deep images of one or both HDFs (Williams et al. 1996; Casertano et al. 2000) to probe the LF to very faint limits that are comparable to the KDF (e.g., Gwyn & Hartwick 1996; Sawicki et al. 1997; Steidel et al. 1999), and those that use wider but shallower ground-based data (e.g., Steidel et al. 1999; Iwata et al. 2003; Ouchi et al. 2004a; Gabasch et al. 2004).

Here we compare our $z \sim 3$ and $\sim$4 LFs with two recent wide but deep surveys, namely, the $z \sim 4$ LF of the Subaru Deep Survey (SDS; Ouchi et al. 2004a) and the $z \sim 3$ and $\sim$4 LFs of the FDF (Gabasch et al. 2004). We also compare our measurement of the LF's faint end with the Steidel et al. (1999) analysis of the HDF-N, for although the HDFs are too small to be truly adequate for LF determination, the Steidel et al. (1999) HDF-based faint-end slope of $\alpha = -1.6$ is often used in the literature, particularly by workers deriving the integrated UV luminosity density and star formation density of the universe at higher redshifts (e.g., Bouwens et al. 2004; Bunker et al. 2004; Dickinson et al. 2004; Giavalisco et al. 2004; Yan & Windhorst 2004).

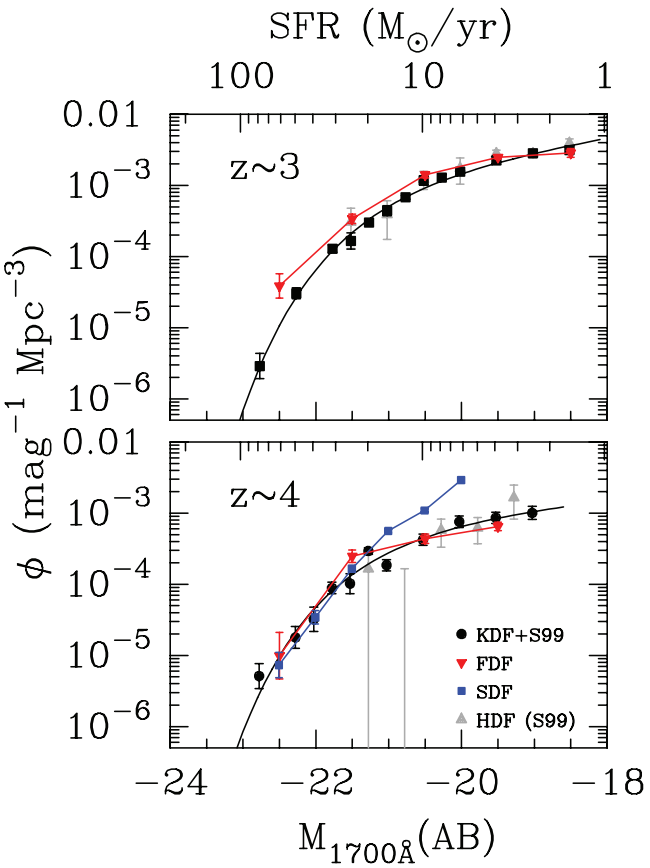

FIG. 9.—Comparisons with other surveys. Rest-frame UV LFs measured in the FDF (Gabasch et al. 2004) at $z \sim 3$ and $\sim$4 and in the SDF (Ouchi et al. 2004a) at $z \sim 4$ are shown. Also shown are the results of the Steidel et al. (1999) analysis of the HDF. The LBG LF is shown as filled circles with error bars that include an estimate of the field-to-field variance determined via bootstrap resampling. The FDF measurement is based on a single small field and does not include an estimate of cosmic variance, while the SDF measurement is based on a single very large field and so is unlikely to be strongly affected by cosmic variance. The LBG LF is in very good agreement with that found in the smaller area FDF at both $z \sim 3$ and $\sim$4. However, while the KDF and the FDF agree with each other, they both disagree with the SDF measurement at $z \sim 4$.

4.4.1. *The Hubble Deep Field LF of Steidel et al. (1999)*

Steidel et al. (1999) have applied a modified version of their color-color LBG selection criteria to the HDF-N in order to extend to fainter magnitudes their $z \sim 3$ and $\sim$4 LF analysis. Their $z \sim 3$ and $\sim$4 HDF LFs are shown in Figure 9 as upward-pointing triangles. At $z \sim 3$, our KDF LF (*filled circles*) is in good qualitative agreement with the Steidel et al. (1999) HDF results, albeit the KDF, with its much larger area and multiple pointings, provides a much more robust and precise measurement. Our faint-end slope $\alpha = -1.43^{+0.17}_{-0.09}$ is significantly shallower than the $\alpha = -1.6$ reported by Steidel et al. (1999) from the HDF (or that by Sawicki et al. [1997], who measured the high-z LFs in the HDF using a different selection technique). As we will discuss in M. Sawicki & D. Thompson (2006a, in preparation), the consequences of this difference in α are not negligible in calculations of the UV luminosity density and star

formation density of the universe. At $z \sim 4$, the HDF data suffer from very poor statistics in addition to potential problems with cosmic variance, but they are lower than the $z \sim 3$ points and, within their large error bars, agree well with the KDF $z \sim 4$ LF.

4.4.2. *The FORS Deep Field*

The FDF reaches depths similar to those of the KDF in a single VLT pointing that covers $\sim$50 arcmin^2, or 29% of the area of the KDF. The fact that it comprises only a single field means that it is highly subject to cosmic variance effects, but, unlike the HDFs, it is sufficiently large that it is not dominated by small number statistics at the faint end. Gabasch et al. (2004) have computed photometric redshifts in the FDF and used them to estimate the galaxy LF at several redshifts and rest-frame wavelengths. They do not present a rest-frame 1700 Å LF, and so we compare our KDF results to their 1500 Å LF. Both of these rest-frame wavelengths probe the part of galaxy SEDs dominated by young, massive stars, and so the systematic biases introduced by this slight difference in wavelengths should not be large.

Figure 9 uses downward-pointing red triangles to show the 1500 Å Gabasch et al. (2004) FDF LFs at $z \sim 3$ and $\sim$4. The agreement between the KDF and FDF is very good, particularly when one considers that the FDF points are based on a single moderately sized field and so do not account for cosmic variance and, moreover, that the FDF and KDF use very different means of sample selection. In particular, we note that both the FDF and the KDF show relatively shallow faint-end slopes at both $z \sim 3$ and $\sim$4 although we feel that the KDF result is much more robust with respect to both cosmic variance (because of its larger area and multiple independent fields) and sample selection (because of the extensive spectroscopic verification of its selection technique).

4.4.3. *The Subaru Deep Survey*

The Subaru Deep Survey (SDS) consists of two large fields (the Subaru Deep Field [SDF] and the Subaru XMM Deep Field [SXDF]). Both fields have been imaged with the Suprimecam large mosaic imager (600 arcmin^2 per field) on Subaru and contain statistically large numbers of galaxies and are very unlikely to be affected by cosmic variance. The SDS $BVRi'$ images allow the selection of $z \sim 4$ and $\sim$5 LBG samples, and Ouchi et al. (2004a) have used such selection to estimate the LBG LFs at these redshifts. A strong limitation of the SDS selection of high-z galaxies is that their color-color selection technique is not calibrated spectroscopically (only a small handful of $z > 1$ spectroscopic redshifts is known in the SDS fields) and must instead rely on models of galaxy colors for its definition. In Figure 9 we use blue squares to plot the $z \sim 4$ rest-frame 1700 Å LF from the SDF field (we omit the SXDF field, which shows virtually identical results but is $\sim$0.5 mag shallower than the SDF). At the bright end, the SDS $z \sim 4$ LF agrees with both the KDF LF (which at these magnitudes is dominated by the Steidel et al. [1999] data) and the FDF result. However, at the faint end, the SDS has a much steeper LF slope than either the KDF or the FDF. The origin of this discrepancy is not immediately clear. One possible explanation is that at faint magnitudes, where as we have seen in § 4.3.3 the scatter in galaxy photometry can strongly affect the estimated effective volumes used in calculating the LF. Ouchi et al. (2004a) do not discuss what galaxy models they used to estimate their $V_{\rm eff}$, but it is possible that while our $V_{\rm eff}$ estimates at $z \sim 4$ are robust to changes in the assumed galaxy SEDs, their effective volumes computed for their BRi' color-color selection are less so. If this is the case, then the $V_{\rm eff}$ in the SDS work may well be underestimated, resulting in overcorrections to the SDS LF at the fainter magnitudes.

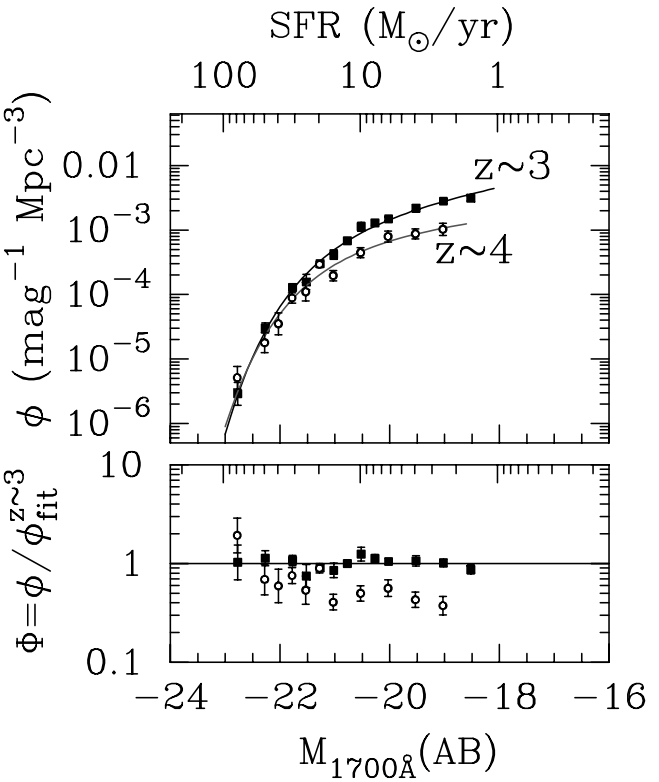

FIG. 10.—Comparison of the $z \sim 4$ and $\sim$3 LFs. The top panel shows the $z \sim 4$ LFs together, with the $z \sim 3$ Schechter fit marked with a thick line. The bottom panel shows the quantity Φ, which measures the fractional deviation of the data (at $z \sim 4$ or 3) from the $z \sim 3$ Schechter fit: perfect agreement between data and the $z \sim 3$ fit would put the points on the horizontal $\Phi = 1$ line.

5. EVOLUTION OF THE LUMINOSITY FUNCTION

In this section we compare our LFs at different redshifts to search for signs of evolution. In what follows, we concentrate on Figures 10 and 11, which compare directly our $z \sim 4$, 3, 2.2, and 1.7 LFs that were determined using the baseline $V_{\rm eff}$ model [100 Myr old starburst with $E(B - V) = 0.15$]. As we elaborate below, our $z \sim 3$ and $\sim$4 LFs are the most unambiguous and free of systematics, whereas the $z \sim 2.2$ and especially $z \sim 1.7$ LFs are likely biased; for this reason in Figure 10 we present the $z \sim 4$ and $\sim$3 LFs only, without the $z \sim 2.2$ and $\sim$1.7 LFs that are shown in Figure 11. Thus, Figure 10 shows only the most solid, bias-free results, while Figure 11 should thus be regarded as presenting all the results, including those biased by systematic effects.

Our LF is most unambiguously constrained at $z \sim 3$, where the combination of KDF and Steidel et al. (1999) data covers the largest range in luminosity, is least dependent on the details of $V_{\rm eff}$ modeling (see § 4.3.3), and puts the tightest bounds on the Schechter function parameters. Consequently, we use the $z \sim 3$ LF as our fiducial reference and compare the other redshift bins to it.

The top panels of Figures 11 and 10 overplot the data and the Schechter function fits, and the bottom panels further highlight the differences between the three redshift samples by showing the data after they have been divided by the Schechter function fit to the $z \sim 3$ LF. Three interesting evolutionary effects can be seen in Figures 11 and 10:

1. A strong increase in the number density of low-luminosity LBGs from $z \sim 4$ to $\sim$3.

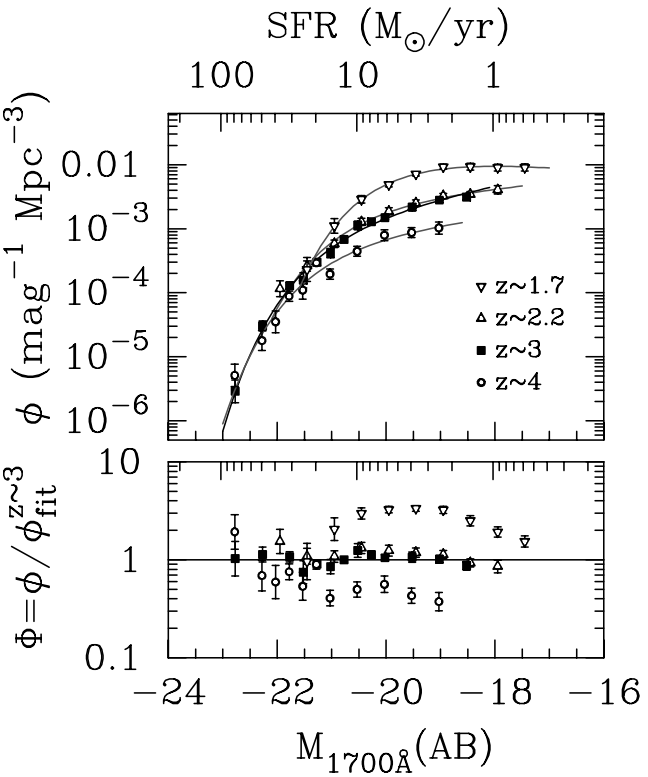

FIG. 11.—Comparison of the $z \sim 4$, 3, 2.2, and 1.7 LFs. As in Fig. 10, but adding the $z \sim 2.2$ and 1.7 LFs. As is discussed in the text, the $z \sim 4$ and $\sim$3 LFs are robust to the details of the analysis procedure, whereas the $z \sim 2.2$ and $\sim$1.7 LFs are less so and are subject to potentially strong systematic uncertainties.

2. An accompanying apparent lack of change in the number density of luminous LBGs from $z \sim 4$ to $\sim$3.

3. A possible continuation of the increase in the number density of low-luminosity galaxies from $z \sim 3$ to $\sim$2.2 and $\sim$1.7.

We discuss these three effects in the following subsections. In each subsection we give a brief phenomenological description of the effect and then concentrate on examining whether the evolution may be explained as an observational effect or whether it instead reflects a true evolution of the underlying galaxy population. Discussion of the possible implications for galaxy evolution of the observed LF evolution is deferred to § 6.

5.1. *Increase in the Number of Low-Luminosity Galaxies from $z \sim 4$ to $\sim$3*

The most striking effect in Figure 10 is the increase in the number density of low-luminosity galaxies from $z \sim 4$ to $\sim$3. For our adopted fiducial $E(B-V) = 0.15$ model, the number of $z \sim 4$ galaxies with $-21 < M_{1700} \lesssim -18$ (i.e., within $M^* \lesssim M_{1700} \lesssim M^* + 2$) is only 0.44 ± 0.05 of the number of such galaxies at $z \sim 3$. In other words, the number of low-luminosity galaxies increases by a factor of 2.3 from $z \sim 4$ to $\sim$3, a result that is statistically significant at the $\sim$11 σ level.

Is this increase a reflection of true galaxy evolution, or is it simply an artifact of some observational bias?

5.1.1. *Can Cosmic Variance be Responsible?*

The effects of large-scale structure can be a problem in small-area surveys, such as the HDFs, or those that consist of a single pointing, such as the FDF. The KDF, however, consists of three widely separated patches on the sky that probe statistically independent parts of the universe. As was discussed in §§ 3 and 4, field-to-field fluctuations are relatively small and the error bars in our LF measurement include a bootstrap estimate of field-to-field variance of our survey. Given that the increase in the number of low-luminosity galaxies from $z \sim 4$ to $\sim$3 is far outside these error bars, we conclude that cosmic variance is unlikely to be responsible for the observed evolution of the faint end of the LF.

5.1.2. *Is Our Modeling of $V_{\rm eff}$ Responsible?*

As we discussed in § 4.3.3, the computation of the LF is dependent on the details of the $V_{\rm eff}$ modeling of the survey. However, as we have shown in § 4.3.3, at $z \sim 4$ and $\sim$3 the change in the LF is remarkably small under a wide range of reasonable assumptions. As is illustrated in the top panel of Figure 7, varying the $V_{\rm eff}$ model assumptions cannot bring the faint ends of the $z \sim 3$ and $\sim$4 LFs into agreement. And as Figure 8 shows further, it appears to be impossible to bring the number density of $\sim M^* + 1$ galaxies at $z \sim 3$ and $\sim$4 into agreement by varying $V_{\rm eff}$ model parameters: even in the most extreme $E(B-V) = 0$ case, where the number densities at $z \sim 4$ and 3 are closest to each other, there remains a very significant deficit of faint galaxies at $z \sim 4$. We therefore conclude that the evolution of the faint end seen in our data is unlikely to be an artifact of the assumptions that underlie our calculation of $V_{\rm eff}$.

5.1.3. *Can Differential Sample Selection be Responsible?*

The color selection criteria used to select galaxies at $z \sim 3$ and $\sim$4 are designed to select galaxies with similar underlying SEDs. However, the color-color selection regions are *not* identical in relation to intrinsic galaxy colors (see Fig. 1), and so we may question whether the deficit of sub-L^* galaxies at $z \sim 4$ results simply because at $z \sim 4$ we are missing galaxies whose SEDs are such that they would have been included in the $z \sim 3$ selection criteria.

If there are significant numbers of such galaxies missing from our sample at $z \sim 4$, then we should be able to include them by expanding the $z \sim 4$ color-color selection region. However, the deficit at the faint end is so large (a factor of 2.3) that no reasonable adjustment to the $z \sim 4$ color selection criteria can remedy it. As Figure 1 shows, the galaxy color model tracks allow the possibility of modifying the $z \sim 4$ color-color selection region by a few tenths of a magnitude. However, we have checked that an increase of even as much as 0.5 mag in both $G - \mathcal{R}$ and $\mathcal{R} - I$ would not be sufficient to bring enough galaxies into the sample to make up the factor of 2.3 deficit (see also Fig. 4 in KDF I). Moreover, even if such a large change to the color selection criteria were permissible, it would result in an automatic increase of the effective volume $V_{\rm eff}$ that would largely counteract any gain from the increase in galaxy numbers.

A further, secondary argument against differential sample selection relies on the fact that the evolution from $z \sim 4$ to $\sim$3 appears to be differential with luminosity (see § 5.2). Since at a given redshift bright and faint LBGs are selected in the same way, a sample selection bias should result in a similar deficit of bright galaxies as of faint ones at $z \sim 4$. The fact that no strong deficit is seen at the bright end of the $z \sim 4$ LF further strengthens the case that the evolution of the faint end is not an artifact of sample selection but is due to real differences between luminous and low-luminosity galaxies.

Overall, we believe that the deficit of faint galaxies at $z \sim 4$ is too large to be accounted for by differences in sample selection and is most likely a real effect that reflects an underlying evolutionary change in the population of sub-L^* galaxies over the $\sim$600 Myr from $z \sim 4$ to $\sim$3. We conclude that the observed

evolution of the faint end of the LF from $z \sim 4$ to ~ 3 is likely real.

5.2. *Luminosity-dependent Evolution from $z \sim 4$ to ~ 3*

Figure 10 suggests that while there is a deficit of $z \sim 4$ sub-L^* galaxies, the number of luminous galaxies remains virtually unchanged from $z \sim 4$ to ~ 3. If real, this differential, luminosity-dependent evolution of the LF hints at important differences in how galaxies of different luminosity evolve at high redshift. We examine the potential implications of this differential evolution in § 6, but first we ask if this differential effect is in fact real.

5.2.1. *Is the Effect Statistically Significant?*

While there can be little doubt of the deficit of faint galaxies at $z \sim 4$, the situation at the bright end is less clear because of the larger uncertainties in the individual data points. In view of this problem, it is tempting to rely on the Schechter functions to compare the bright-end LFs, and indeed the $z \sim 3$ and ~ 4 Schechter function fits are in excellent agreement at the bright end. However, such a comparison can be misleading because the Schechter function fits incorporate data at all luminosities, and so at the bright end the fits may well be biased by the weight of the faint-end data where the statistics are so much better. It would be more robust to compare the $z \sim 3$ and ~ 4 number densities directly, but this is not straightforward because the data at $z \sim 4$ and ~ 3 are sampled differently and cover somewhat different magnitude ranges in the two redshift bins.

To overcome these difficulties, we instead compare how the data at $z \sim 3$ and ~ 4 deviate from our $z \sim 3$ Schechter function fit. We proceed as follows. First, we calculate the quantity $\Phi(M)$, which is the ratio between the data and the $z \sim 3$ Schechter function fit,

$$\Phi(M) = \phi_{\rm data}(M)/\phi_{\rm fit}^{z\sim 3}(M). \quad (12)$$

The bottom panels of Figures 11 and 10 show $\Phi(M)$ for our four redshift samples. Note that $\Phi(M)$ is computed for each of the four redshift samples, but in each case the data are divided by the same, $z \sim 3$, Schechter function. Consequently, $\Phi(M)$ is close to 1 at $z \sim 3$ (reflecting the fact that the $z \sim 3$ Schechter function is a good representation of the $z \sim 3$ data) but deviates from 1 for the other redshift bins, and in particular for the $z \sim 4$ sample. We then compute the average $\Phi(M)$, namely, $\bar{\Phi}$, for galaxies brighter and fainter than $M_{1700} = -21.0$ (i.e., $\sim M^*$) at both redshifts of interest here, $z \sim 3$ and ~ 4. Finally, the ratio of the $\bar{\Phi}$ at the two redshifts then tells us the amount of number density evolution that the given population undergoes. By comparing the ratios of the $\bar{\Phi}$, we effectively cancel out the dependence of our comparison on the $z \sim 3$ Schechter function fit and are comparing the data at $z \sim 3$ and ~ 4 directly.

Figure 12 shows these $\bar{\Phi}$ ratios for the bright (horizontal axis) and faint (vertical axis) ends of the LF. The quantities in Figure 12 are always shown as evolution with respect to the $z \sim 3$ case (i.e., we plot $\bar{\Phi}_z/\bar{\Phi}_{z\sim 3}$). Three evolutionary scenarios are marked for reference: locations on the vertical straight line indicate no number density change in the bright end of the LF, locations on the horizontal solid line indicate no evolution in the faint end, and locations on the diagonal line indicate equal number density evolution at the bright and faint end. The intersection of the three lines at (1, 1) marks the case of a nonevolving LF.

As Figure 12 shows, there is substantial change in the number density ratio of sub-L^* galaxies from $z \sim 4$ to ~ 3: $\bar{\Phi}(z \sim 4)/\bar{\Phi}(z \sim 3) = 0.44 \pm 0.05$ for faint galaxies, indicating that there is a 2.3-fold increase in the number density of faint galaxies that is statistically significant at the $\sim 11\ \sigma$ level (this is the faint-end evolution we discussed in § 5.1). At the same time, however, Figure 12 also shows that the number density of luminous galaxies at $z \sim 4$ is virtually unchanged with $\bar{\Phi}(z \sim 4)/\bar{\Phi}(z \sim 3) = 0.87 \pm 0.19$.

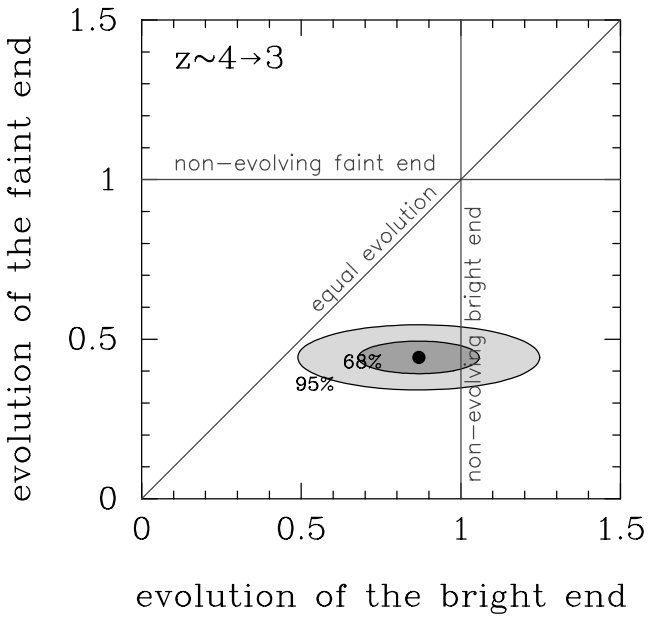

FIG. 12.—Luminosity-dependent evolution of the number density of galaxies. The horizontal axis shows the amount of change in the density of luminous galaxies ($M_{1700} < -21$) parameterized as the LF as the ratio between the average normalized densities $\bar{\Phi}$ at $z \sim 4$ and ~ 3. The vertical axis shows the same quantity but for low-luminosity galaxies, $M_{1700} > -21$. The straight lines represent three fiducial cases: no faint-end evolution (*horizontal line*), no bright-end evolution (*vertical line*), and equal evolution at the bright and faint ends (*diagonal line*). The filled circle and ellipses represent the amount of evolution from $z \sim 4$ to ~ 3 and the associated 68% and 95% confidence regions. There is substantial evolution of the low-luminosity population from $z \sim 4$ to ~ 3 that is statistically significant at the 98.7% level.

To properly test whether the evolution of the LF is differential with luminosity, we must of course consider the joint uncertainty for the bright- and faint-end cases. The joint 1 and 2 σ uncertainties are illustrated by the error ellipses in Figure 12, and in this context, the distance of the $z \sim 4 \rightarrow 3$ point from the diagonal "equal evolution" line indicates the amount of differential evolution. We find that 98.5% of random realizations of the data in Figure 12 are above the diagonal "equal evolution" line and so inconsistent with the differential evolution scenario. This strongly suggests that the high-z galaxy population is undergoing differential, luminosity-dependent evolution. Stronger confirmation of this assertion will require improved constraints on the bright end of the LF; this confirmation will require LBG surveys with areas of several square degrees, i.e., an order of magnitude greater than used in the work of Steidel et al. (1999) that provides the bulk of the statistics at the bright end of the LF here.

5.2.2. *Can Differential Sample Selection be Responsible?*

The bright and faint galaxies in a given redshift bin are selected using identical selection criteria. Moreover, given that the bright-end galaxies come primarily from the Steidel et al. (1999) work and the faint-end ones exclusively from the significantly deeper KDF, they suffer from similar amounts of photometric scattering out of the color-color selection regions (at any rate, this scatter is accounted for through the $V_{\rm eff}$ approach). Consequently, as long as the bright and faint samples within that bin

consist of the same mix of galaxy SEDs, any systematic effect should be reflected in the same way at both the bright and faint ends.

At the bright end, the LF is constrained mainly by the Steidel et al. (1999) results, while the faint end is dominated by the KDF data. It is possible that despite our great care some subtle, unknown selection effect remains between the two samples. However, we consider this possibility to be extremely unlikely in view of the virtually identical observational approaches and sample selection and analysis techniques that were applied to the two data sets.

Finally, we cannot rule out the possibility that the faint and bright galaxies at a given redshift represent populations that differ not just in luminosity but also in some other property such as, for example, the amount of interstellar dust. However, if such true, intrinsic luminosity-dependent differences do exist, they do not alter our conclusion that the LF evolution is differential with luminosity but merely shift the focus to a more specific, but still real, evolutionary effect. We conclude that while the different composition of the samples with luminosity within a given redshift bin may explain the differential, luminosity-dependent evolution we see from $z \sim 4$ to ~ 3, such differences in intrinsic galaxy properties would only underscore the point that the luminous and faint UV-selected galaxy subpopulations are different and not simply scaled analogs of each other.

5.3. *Evolution from $z \sim 3$ to ~ 2.2*

At $z \sim 2.2$ (and ~ 1.7) we do not have sufficient statistics in the KDF data to say much about the LF for galaxies brighter than L^*, nor, unlike at $z \sim 3$ and ~ 4, are there published results from the Steidel group that we could use to extend our luminosity range. Consequently, we focus on the evolution of the faint end of the LF only.

Figure 11 shows evolution of the LF from $z \sim 3$ to ~ 2.2 (see also Figs. 4 and 7). The LF appears to undergo only a small amount of evolution from $z \sim 3$ to ~ 2.2 for our fiducial $V_{\rm eff}$ model [100 Myr old starburst with $E(B-V) = 0.15$]: there is only a factor of 1.1 ± 0.1 increase in the number density of sub-L^* galaxies with cosmic time. However, as Figure 8 illustrates, our chosen fiducial $V_{\rm eff}$ model happens to give nearly the minimal evolution in the LF from $z \sim 3$ to ~ 2.2, and applying $z \sim 2.2$ $V_{\rm eff}$ models with either larger or smaller extinction values results in stronger LF evolution. For example, using the $E(B-V) = 0.3$ model for both $z \sim 3$ and ~ 2.2 gives a factor of ~ 1.2 increase in the number density of sub-L^* galaxies, and a yet stronger evolution (factor of 2.3 number density increase) results with the $E(B-V) = 0$ model. Likewise, adopting a $V_{\rm eff}$ model that assumes a galaxy population with a range of $E(B-V)$ values gives evolution that is stronger than our fiducial case: using a flat number distribution over $E(B-V) = 0-0.3$ results in an increase of 1.24 ± 0.08 in the number density of galaxies fainter than $M_{1700} = -21.0$.

The above examples illustrate that the measurement of the $z \sim 2.2$ LF has a systematic uncertainty that depends on our assumptions about the properties of the UV-bright galaxy population at these redshifts. Whereas the $z \sim 4$ and, especially, $z \sim 3$ LFs are only weakly dependent on the assumed $V_{\rm eff}$ model, the $z \sim 2.2$ LF shows a much stronger dependence. A direct result of this dependence is that we cannot unambiguously determine the $z \sim 2.2$ LF and the amount of LF evolution from $z \sim 3$ to ~ 2.2.

Despite these limitations, we can still put some useful constraints on the evolution of the LF over the 800 Myr from $z \sim 3$ to ~ 2.2. As Figure 7 shows (see also Fig. 8), there are at least as many galaxies at a given UV luminosity at $z \sim 2.2$ as there are at $z \sim 3$. Depending on the adopted $V_{\rm eff}$ models, the number density can be close to identical at the two redshifts but potentially can be up to a factor of ~ 2 higher at $z \sim 2.2$ than at $z \sim 3$. Note that this conclusion could potentially be further affected by differential sample selection between $z \sim 3$ and ~ 2.2, although it is unlikely that sample selection differences will drastically modify the results given that the $z \sim 3$ and ~ 2.2 color-color selection criteria select galaxies with similar intrinsic SEDs (Fig. 1) and are likely to capture the bulk of UV-bright galaxies at these redshifts, as can be seen in Figure 5 of KDF I.

In summary, we can conclude that there are probably at least as many, and possibly more, sub-L^* UV-bright galaxies at $z \sim 2.2$ as there are at $z \sim 3$. There is no evidence for a decline in galaxy numbers with cosmic time.

5.4. *Toward $z \sim 1.7$?*

Figure 11 shows apparent strong evolution of the LF from higher redshifts down to $z \sim 1.7$. However, we feel that the LF measurement at $z \sim 1.7$ is not reliable given systematic uncertainties in our LF estimation at this redshift. The main systematic problem stems from the lack of robustness in the modeling of $V_{\rm eff}$ for the $z \sim 1.7$ sample. While we have high confidence that our $V_{\rm eff}$ modeling is robust at $z \sim 4$ and ~ 3 and somewhat less accurate but still partially reliable at $z \sim 2.2$, we have far less confidence in its reliability at $z \sim 1.7$, as explained below.

While our $V_{\rm eff}$ modeling for the higher redshift bins accurately reproduces the observed redshift distribution of galaxies at $z \sim 4$, 3, and 2.2, it fails to do so at $z \sim 1.7$, as can be seen in Figure 2. Another manifestation of this problem can be seen in Figure 1, where the model colors of $z = 1.7$ galaxies ($z = 1.7$ is the mean redshift of the observed spectroscopic sample) are too red in $U_n - G$ compared to the color-color selection region defined by Steidel et al. (2004). The upshot of these mismatches is that we do not have confidence that our modeling of $V_{\rm eff}$ accurately reflects the true volume from which the color-color–selected galaxies in the $z \sim 1.7$ sample are drawn. An inaccurate calculation of the survey volume has the potential to strongly affect the computed LF. If, for example, our $V_{\rm eff}$ modeling underestimates the volume of the survey, then this will translate into an overestimate of the number density of galaxies. This may well be happening and would explain the high-density normalization of the $z \sim 1.7$ LF seen in Figure 11.

There are several possible underlying reasons why our $V_{\rm eff}$ modeling at $z \sim 1.7$ may be inaccurate. We do not at present know what $E(B-V)$ value or star formation history is appropriate for $z \sim 1.7$ galaxies. At $z \sim 3$, good estimates of these quantities are known from observations (see § 3.1.1), and it is not unreasonable that similar $E(B-V)$ and age values are also applicable at $z \sim 4$ and ~ 2.2 given that these redshift bins are only 0.6 and 0.8 Gyr away from $z \sim 3$. By $z \sim 1.7$, however, 1.8 Gyr after $z \sim 3$, the galaxy population may have significantly different properties than it does at $z \sim 3$ as galaxies evolve toward the more prosaic, less starburst-dominated and less dust-obscured galaxy population observed by $z \sim 1$. Assuming the $z \sim 3$ reddening and starburst age values for the $z \sim 1.7$ population may well strongly skew our $V_{\rm eff}$ estimate and be reflected in the mismatches seen at $z \sim 1.7$ in Figures 1 and 2.

In addition to systematic problems with the $V_{\rm eff}$ calculation, a second potential systematic effect may be affecting our LF measurement at $z \sim 1.7$: it is possible that our $z \sim 1.7$ color-color–selected sample may be significantly contaminated by low-z galaxies. While the spectroscopically determined contamination rate in the $\mathcal{R} \sim 24-25.5$ $z \sim 1.7$ sample of Steidel et al. (2004) is less than 5%, even small systematic offsets in the color-scale

calibration between our and their photometric systems can result in a drastically increased contamination fraction. We stress here that our modeling of $V_{\rm eff}$ accounts for the scatter of high-z galaxies between the different high-z color-color selection regions but does not account for the scatter of low-redshift ($z \lesssim 1$) galaxies into the high-z selection boxes. At $z \sim 4$, 3, and 2.2, the scatter from low redshift is unlikely to be an issue given that these selection regions are far removed from the $z < 1$ loci in color-color space. In contrast, the $z \sim 1.7$ selection region is in close proximity to the region of space dominated by $z \lesssim 1$ galaxies, and it is possible that the low-z interloper contamination rate may be quite high. A high foreground contamination fraction would result directly in an overestimate of the number density of $z \sim 1.7$ galaxies and an inaccurate LF.

In summary, unlike at $z \sim 4$, 3, and 2.2., we are not confident in our determination of the LF at $z \sim 1.7$. We have chosen to present the $z \sim 1.7$ LF here for completeness, but we urge the reader to regard it with much caution.

6. DISCUSSION: THE EVOLVING GALAXY POPULATION

6.1. *Summary of the Observational Evidence*

To summarize § 5, the number of faint (sub-L^*) galaxies increases with cosmic time, with ∼2.3 times more galaxies of a given luminosity at $z \sim 3$ than at $z \sim 4$; this increase is statistically significant at the 11 σ level and, as we have argued in § 5.1, it seems very unlikely that it is an artifact of differential sample selection or cosmic variance in our data. At the same time, the evolution from $z \sim 4$ to ∼3 appears to be differential with luminosity, since while the faint end of the LF evolves significantly the bright end appears to remain virtually unchanged. This differential effect is statistically significant at the 97% level and is unlikely to be due to a systematic bias. The case for evolution from $z \sim 3$ to lower redshifts is less clear because of potential systematic biases, although it is possible that the number density of sub-L^* galaxies continues to increase to at least $z \sim 2.2$; our analysis suggests that it is extremely unlikely that the number density of sub-L^* galaxies at $z \sim 2.2$ is lower than at $z \sim 3$.

Because our results at $z \sim 3$ and ∼4 are highly robust (while those at lower redshifts are less so), in what follows we focus on the evolution of these two epochs.

6.2. *The Importance of Luminosity Function Evolution*

The evolution of the LF measures only the evolution of the galaxy population as a whole and does not necessarily imply a direct correspondence in the evolution of individual galaxies. The observed evolution of the LF's faint end can be interpreted equally well as a change in the number density or luminosity of the observable population. However, individual galaxies are free to change their luminosities following trajectories that are far more complicated than a direct increase in luminosity or number density as the evolution of the faint end of the LF might naively suggest. Likewise, the apparent constancy of the LF's bright end does not necessarily imply that the luminous galaxies that populate it do not themselves evolve. Evolution of the LF is clearly not a direct probe of the evolution of individual galaxies.

However, while the luminosity or number density evolution of the LF does not necessarily reflect a direct corresponding evolution in the properties of individual galaxies, the fact that the LF does evolve constitutes an important suggestion that its constituent individual galaxies do evolve over time. Furthermore, the fact that the LF's evolution appears to be luminosity dependent suggests that the evolution of individual galaxies is also differential with luminosity. It suggests that there may be real physical differences between low- and high-luminosity systems in properties such as the supply of gas available for star formation, the merger rates that may trigger such star formation, properties of the dust that obscures it, or the effectiveness of feedback that can regulate it.

6.3. *Some Evolutionary Speculations*

Both semianalytic and smoothed particle hydrodynamics galaxy evolution models have been used to predict the shape of the LBG LF at different redshifts (e.g., Somerville et al. 2001; Nagamine et al. 2004). However, such models tend to produce a relatively constant, unevolving LF: we are not aware of any predictions in the literature for the evolution of the sub-L^* end of the LF that we observe from $z \sim 4$ to ∼3. It would be interesting to see what modifications to these sophisticated galaxy formation models can reproduce a differentially evolving LF. In the meantime, in the absence of such predictions, we turn to some simple phenomenological speculations about the possible nature of the evolution of individual galaxies that underlies the observed evolution in the LF. We explore three heuristic evolutionary scenarios that we use to illustrate how varying some simple properties of individual galaxies can mirror the observed differential, luminosity-dependent evolution of the LF.

Our three heuristic models are motivated as follows. SED studies suggest that high-z galaxies likely undergo episodes of intense star formation followed by more quiescent periods, and it is plausible that such episodes occur several times in the life of a high-z galaxy (see, e.g., Sawicki & Yee 1998; Shapley et al. 2001). Moreover, it is clear that UV-selected high-z galaxies are obscured by significant amounts of starlight-absorbing dust (e.g., Meurer et al. 1997; Sawicki & Yee 1998; Ouchi et al. 1999; Shapley et al. 2001; Papovich et al. 2001; Vijh et al. 2003), and there is no reason to believe that the properties of this dust, such as its opacity or large-scale geometry, remain constant with time or SFR. Two of our heuristic scenarios (scenarios B and C) are thus motivated by the possibility that the properties of dust or of the starbursting episodes evolve with time and/or luminosity. Meanwhile, scenario A investigates the more simple picture that in the evolution of the faint end we are seeing the very first appearance of many of the low-luminosity galaxies.

We stress that our three heuristic scenarios are not meant as a comprehensive survey of all possible evolutionary mechanisms. Clearly there are many others, but we focus on these three to illustrate some interesting possibilities and motivate future follow-up studies.

6.3.1. *Scenario A: The First Appearance of Low-Luminosity Galaxies?*

One of the simplest possible pictures of LBGs is that they are objects that form their stars at a constant, unvarying rate for long periods of time. SED modeling can be used to constrain the ages of the ongoing episodes of star formation in LBGs, and while such ages are notoriously dependent on the assumed star formation histories, it is the constant SFR scenarios that yield the oldest ages (Sawicki & Yee 1998; Papovich et al. 2001), thus providing upper bounds. Shapley et al. (2001) have modeled a large sample of relatively luminous (typically $\mathcal{R} \sim 24$) $z \sim 3$ LBGs under the assumption of constant star formation. Of 72 luminous $z \sim 3$ galaxies in their analysis, only 38% have ages older than 0.6 Gyr, and so only 38% of the luminous LBGs seen at $z \sim 3$ were forming stars at $z \sim 4$; the remainder must have begun their current episodes of star formation more recently than

$z \sim 4$. Meanwhile, since the number density of luminous LBGs is the same at $z \sim 4$ as at $z \sim 3$, 62% of luminous $z \sim 4$ LBGs must have ceased star formation before $z \sim 3$ to be replaced by the younger starbursts. Similar reasoning applied by Iwata et al. (2003) to an even earlier epoch suggests that only 20% of luminous LBGs at $z \sim 5$ can still be seen at $z \sim 3$, while the remaining 80% must have been replaced to keep the bright end of the LF constant. These arguments suggest that individual LBGs cannot be in a steady star-forming state but at best are in a quasi–steady state where individual galaxies fade in and out of a given magnitude bin to keep the number density constant, even if timescales for such fading are long. Indeed, a detailed analysis by Ferguson et al. (2002) calculated that the star formation histories of (luminous) $z \sim 3$ LBGs are inconsistent with the observed number density of these galaxies at $z \sim 4$ unless episodic bursts of star formation are invoked.

However, the situation may well be different for low-luminosity galaxies: there it is possible to reproduce the observed LF evolution by postulating that a low-luminosity LBG remains at a constant UV luminosity once it starts forming stars. The age distribution of sub-L^* LBGs has not yet been measured, but let us assume that, as for the luminous LBGs, only 38% of low-luminosity $z \sim 3$ LBGs are old enough to have been present at $z \sim 4$. If at the same time we assume that none of the $z \sim 4$ LBGs have ceased star formation to fade out of the sample, then we can reproduce the $\sim$2.3-fold number density evolution of low-luminosity LBGs by assuming that new low-luminosity LBGs are being simply added to the population between $z \sim 4$ and $\sim$3.

Under this scenario it is possible that we are seeing large numbers of low-luminosity LBGs "light up" for the first time in the epoch between $z \sim 4$ and $\sim$3. If this is the case, then we may expect them to have low metallicities, a property that should allow us to test this scenario through future observations. This picture is perhaps somewhat akin to "downsizing" scenarios that postulate that star formation activity shifts to lower luminosity objects over time. It is possible, however, to explain the observed LF evolution with other evolutionary models, as we discuss next.

6.3.2. *Scenario B: Evolution in the Properties of Star-forming Bursts?*

The LF evolution may also (or instead) be related to the frequencies, durations, or intensities of the starbursting episodes that likely rule high-z galaxies. While the constant star formation assumption in SED fitting of LBGs can result in relatively old starburst ages, other assumed star formation histories can yield significantly shorter star formation episodes (Sawicki & Yee 1998; Papovich et al. 2001), and it is plausible that such periods of relatively brief elevated star formation may reoccur several times in each galaxy between $z \sim 4$ and $\sim$3. Such fluctuating SFRs are a feature of some models of LBGs (e.g., Nagamine et al. 2004). As the SFR in a galaxy fluctuates over time, that galaxy will move back and forth between magnitude bins in the LF. If the characteristic intensities, durations, or frequencies (duty cycles) of these episodes decrease with redshift, the resulting effect will be to alter the shape of the LF.

In particular, if intrinsically low-luminosity galaxies are spending progressively more time in the state of elevated star formation (be it because the intense episodes are longer or occur more frequently), then the faint end of the LF will steepen; an increase in the intensities (i.e., SFRs) of the starburst episodes will have a similar effect on the LF. At the same time, the bright end of the LF will remain constant if the characteristics of the starbursting episodes in the intrinsically luminous galaxies remain fixed with time.

It is not obvious what mechanism could be responsible for the change in the duration or intensity of star formation, but one possibility is that star formation becomes more robust against self-disruption by feedback effects as their host dark matter halos accrete material with time. The lack of evolution at the bright end would then suggest that the mechanism depends on halo mass and has saturated for the more massive galaxies so that it cannot evolve further with time even though the dark matter halos themselves may still be growing.

6.3.3. *Scenario C: Evolution in the Properties of Dust?*

The third scenario we examine is linked to the possible evolution in the properties of interstellar dust in high-z galaxies. Even a small change in the amount of dust can have a strong effect on the observed UV luminosity of a galaxy while leaving its colors relatively unaffected. For example, a decrease in extinction from $E(B - V) = 0.25$ to 0.1 in a $z \sim 3.5$ LBG would produce a sixfold increase in its M_{1700}, enough to match the evolution of the faint end of the LF we see between $z \sim 4$ and $\sim$3, and yet would result in galaxy colors that still remain well within the LBG color-color selection criteria. Because the luminosity change of the bright end of the LF is very strongly ruled out by the data, such evolution in dust properties would have to be differential with luminosity, which suggests that such changes cannot be due to differences in sample selection between redshift bins but could be due to real changes in dust properties.

Again, it is not immediately clear what mechanism could result in a change in effective dust opacity in low-luminosity but not high-luminosity LBGs. If the dust evolution scenario is correct, then it may reflect time-dependent changes in the properties of dust grains in the sub-L^* LBGs or in the amount of obscured versus unobscured area visible in each LBG.

While it is interesting to speculate about the nature of the underlying evolution of individual galaxies that is reflected in the evolving LF, clearly, the LF by itself is insufficient to discriminate between the possible mechanisms that are responsible. To understand what drives the changes we see, we will have to turn to follow-up studies that compare the properties of dust and star formation in high-z galaxies as a function of redshift and luminosity.

6.3.4. *The Way Forward*

A key result is that we have identified luminosity and redshift as important variables in galaxy evolution at high redshift. We can use this fact to seek the nature of the underlying evolutionary mechanism by comparing diagnostics of dust, age, etc., as a function of L and z. While LBG follow-up studies to date have primarily focused on luminous galaxies at $z \sim 3$, now that we know that galaxy evolution depends on L and z, extending such studies as a function of luminosity and redshift provides an attractive way to gain valuable insights into how galaxies form and evolve.

As we have illustrated in § 6.3, evolution in the properties of starbursting episodes or in the amount or distribution of interstellar dust may underlie the evolution of the LF. One line of attack then is to compare the broadband SEDs of LBGs as a function of L and z. SED studies have already yielded insights into the extinction, starburst ages, and stellar masses of LBGs (e.g., Sawicki & Yee 1998; Papovich et al. 2001; Shapley et al. 2001), but by applying these SED fitting techniques to different LBG subpopulations, we can look for systematic trends that may reflect the dominant evolutionary mechanisms. Evolution in the rest-frame UV–optical SEDs of luminous ($L \gtrsim L^*$) LBGs from $z \sim 4$ to $\sim$3 suggests buildup of stellar mass and possibly a finely

tuned interplay between increasing dust content and SFRs (Papovich et al. 2004) even in the observed absence of evolution in the LF; it will be interesting to see what SED analyses tell us about the evolution of sub-L^* LBGs.

Another approach will be to compare the detailed spectra of LBGs as a function of z and L. A composite spectrum representing a $\mathcal{R} \sim 24.5$ LBG has yielded detailed insights into the properties of LBG stellar populations, outflows, etc. (Shapley et al. 2003). Comparing composite spectra of LBGs of different luminosity and at different epochs may yield key insights into what makes LBGs different as a function of L and z.

Yet another line of attack is to measure galaxy clustering as a function of both luminosity and redshift as this measurement will let us relate the potentially time-varying UV luminosity to the more stable dark matter halo mass; while studies of the luminosity dependence of clustering have been attempted in the past (Giavalisco & Dickinson 2001; Ouchi et al. 2004b), they have relied on either very small fields or spectroscopically untested selection techniques; the KDF is designed specifically with studying the dependence of clustering on LBG luminosity in mind, and we will attack this issue in M. Sawicki & D. Thompson (2006b, in preparation).

There are two very important advantages that such future differential studies will have. Foremost, (1) from an experimenter's viewpoint, we now know that L and z are variables that affect how galaxies evolve. We can thus be confident that "varying" L and z will yield a "response" in galaxy properties linked to evolutionary mechanisms and that the lack of such response will equally importantly rule out a candidate evolutionary mechanism. At the same time, (2) while such studies will likely build on previously developed techniques as illustrated above, by comparing results as a function of L and z they will use these techniques in a an essentially differential sense, thereby reducing our current reliance on theoretical models or low-z analogs. Differential measurements are always much easier and more robust than absolute ones, making such differential studies extremely attractive.

We feel that important insights lie ahead using this differential approach, and we will pursue such studies in the near future.

7. SUMMARY AND CONCLUSIONS

In this paper we have used our very deep KDF catalogs of $U_nG\mathcal{R}I$ color-color–selected galaxies at high redshift to construct the LFs of UV-selected galaxies at $z \sim 4$, 3, 2.2, and 1.7. As we discuss in detail in KDF I, these catalogs use the very same color-color selection techniques as are used by Steidel et al. (1999, 2003, 2004) to select their galaxy samples at these redshifts. Moreover, we use the same effective volume ($V_{\rm eff}$) approach to computing the galaxy LF as used by Steidel et al. (1999) at $z \sim 3$ and $\sim$4. However, our KDF data select galaxies to $\mathcal{R}_{\rm lim} = 27$, a magnitude and a half deeper than that previous work, and so allow us to probe the faint, sub-L^* end of the galaxy LFs at these and lower redshifts. Our analysis probes the population to galaxies that have SFRs of $\sim 1\ M_\odot\ {\rm yr}^{-1}$ in the absence of interstellar dust and in our assumed $(\Omega_M, \Omega_\Lambda, H_0) = (0.3, 0.7, 70\ {\rm km\ s^{-1}\ Mpc^{-1}})$ cosmology.

Spectroscopic redshifts for a large sample of galaxies to $\mathcal{R}_{\rm lim} \sim 27$ would be observationally extremely expensive, and so the estimate of the LF at these faint limits must at present rely on photometric redshifts or its cousin color-color selection. Several attempts to estimate the faint end of the LF at these redshifts have been made in the past. However, ours has the important and unique combination of using a color-color selection technique that is well understood and well tested spectroscopically while at the same time drawing on a galaxy sample that is taken from a large area of 169 arcmin^2, thus giving good statistics, and spanning three spatially independent fields, thus allowing us to control for cosmic variance. To our knowledge no other survey in existence has this combination of favorable and important characteristics. We thus believe that ours is the most reliable estimate of the faint end of the LF at these redshifts to date.

We have carried out detailed studies to understand the potential systematic biases in our LF analysis. We find that field-to-field variance or uncertainties due to k-corrections do not significantly affect our results and find instead that the largest source of systematic uncertainty lies in the estimate of the effective survey volume, $V_{\rm eff}$. We found that our results are robust to how we estimate $V_{\rm eff}$ at $z \sim 3$ and $\sim$4; however, the estimate of $V_{\rm eff}$ introduces a source of systematic uncertainty into the LF at $z \sim 2.2$ and $\sim$1.7. In addition, we also suspect that contamination by low-z interlopers may be an additional source of uncertainty at $z \sim 1.7$. Overall, we are highly confident of our LF estimates at $z \sim 4$ and $\sim$3, we feel we can use the $z \sim 2.2$ results to place limits on the shape of the LF, and we are not confident of the $z \sim 1.7$ estimate.

In light of the preceding discussion, the results of our analysis can be summarized as follows:

1. The faint-end slope of the LBG LF at $z \sim 3$ and $\sim$4 is shallower than the $\alpha = -1.6 \pm 0.13$ previously reported at $z \sim 3$ by Steidel et al. (1999) using HDF data. We find $\alpha = -1.43^{+0.17}_{-0.09}$ at $z \sim 3$ and $-1.26^{+0.40}_{-0.36}$ at $z \sim 4$. While formally consistent with the Steidel et al. (1999) α, our more accurate, lower α may force a factor of 2 downward readjustment of many of the recent UV-based estimates of the density of star formation in the universe at $z \sim 3$ and above, an issue that we address in M. Sawicki & D. Thompson (2006a, in preparation).

2. We find strong evolution in the number density of faint (sub-L^*) LBGs over the 600 Myr from $z \sim 4$ to $\sim$3: there are 2.3 times more sub-L^* LBGs at $z \sim 3$ than at $z \sim 4$. This result is statistically secure at the 11 σ level, and we believe it to be independent of systematic biases due to cosmic variance, sample selection differences, surface brightness selection differences, assumptions about the cosmological model, or differential k-correction effects.

3. While the faint end of the LF evolves from $z \sim 4$ to $\sim$3, the bright end appears to remain unchanged. This differential, luminosity-dependent evolution is statistically significant at the 98.5% level, where the limitation in our confidence comes from the small number statistics of the bright end of the LF. An improvement in the level of confidence here will require analysis of $z \sim 4$ and $\sim$3 LBG surveys that are significantly larger than even the largest that have been studied to date. If the differential evolution is real, then it may allow new, differential approaches to the study of LBGs.

4. It is not clear whether the evolution of the faint end of the LF continues to lower redshift because of potential systematic biases at $z \sim 2.2$ and $\sim$1.7. We find that our estimate of the $z \sim 2.2$ LF depends on our assumptions about the makeup of the galaxy sample at this redshift. Despite these systematic uncertainties, we can nevertheless conclude that there are at least as many sub-L^* galaxies at $z \sim 2.2$ as there are at $z \sim 3$.

5. At $z \sim 1.7$, systematic effects make it difficult to be confident of the reliability of our LF determination at that redshift.

The two most intriguing results of the work presented here are the increase in the number density of sub-L^* galaxies from $z \sim 4$ to $\sim$3 and the possibility of differential, luminosity-dependent evolution over that redshift interval. As we discussed in §§ 4.3

and 5.1, the increase in the number of low-luminosity galaxies is a robust result that is both statistically highly significant and at the same time unlikely to be an artifact of some systematic bias. The presence of differential evolution of the galaxy population is a less secure result (98.5% statistical probability; § 5.2) that will need to be confirmed with much larger, shallow LBG surveys. However, we note that differential LF evolution is not unexpected given that there is no reason to think that galaxies across a range of UV luminosity (and so, to first order, a range of different SFRs) are straightforwardly scaled analogs of each other.

The evolution of the faint end of the population raises the intriguing question, what processes in individual galaxies underlie the observed evolution of the population? A wide range of possible evolutionary mechanisms may be at play, ranging from changes in the properties of starbursting episodes that seem to occur in these galaxies to evolution in the amount or properties of interstellar dust. Discerning what mechanism is responsible will be important for our understanding of how high-redshift galaxies form and evolve.

One avenue of attack on this problem is suggested by the fact that the evolution of the galaxy population may be differential with luminosity. If the population does evolve differentially with luminosity, then comparing the properties of LBGs as a function of luminosity and (for faint LBGs) of redshift may point us toward the underlying mechanism. If, as some studies suggest (e.g., Giavalisco & Dickinson 2001), UV luminosity is a tracer of the dark matter halo mass, then we will first be able to link the differential evolution of the population to a mass scale. We are starting to pursue this line of attack with our KDF data (M. Sawicki & D. Thompson 2006b, in preparation). Another set of insights will be possible from comparing the SEDs of LBGs or the details of their composite spectra; such studies have to date been focused on relatively luminous LBGs at $z \sim 3$, where they have yielded insights into, e.g., the starburst ages, extinction, and the state of their interstellar media (Sawicki & Yee 1998; Papovich et al. 2001; Shapley et al. 2001, 2003). The intriguing possibility of luminosity-dependent evolution of the LBG population opens up the possibility for such studies in a way that is differential and so largely independent of the systematics associated with using models or low-redshift analogs. We will explore such approaches in upcoming work.

We thank the Palomar time allocation committee for a generous time allocation that made this project possible and the staff of the W. M. Keck Observatory for their help in obtaining these data. We are grateful to Chuck Steidel for his encouragement and support of this project and Jerzy Sawicki for many useful comments. We also thank Masami Ouchi and Armin Gabasch for providing their LF data points in tabular format and Naveen Reddy and Ikuru Iwata for useful discussions. Finally, we 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.

REFERENCES

Adelberger, K. L., Steidel, C. C., Giavalisco, M., Dickinson, M., Pettini, M., & Kellogg, M. 1998, ApJ, 505, 18
Adelberger, K. L., Steidel, C. C., Pettini, M., Shapley, A. E., Reddy, N., & Erb, D. K. 2005, ApJ, 619, 697
Adelberger, K. L., Steidel, C. C., Shapley, A. E., & Pettini, M. 2003, ApJ, 584, 45
Ando, M., Ohta, K., Iwata, I., Watanabe, C., Tamura, N., Akiyama, M., & Aoki, K. 2004, ApJ, 610, 635
Arnouts, S., et al. 2005, ApJ, 619, L43
Barger, A. J., et al. 1998, Nature, 394, 248
Blain, A. W., Smail, I., Ivison, R. J., & Kneib, J.-P. 1999, MNRAS, 302, 632
Bouwens, R. J., Illingworth, G. D., Blakeslee, J. P., & Franx, M. 2006, ApJ, in press (astro-ph/0509641)
Bouwens, R. J., et al. 2004, ApJ, 606, L25
Bruzual A., G., & Charlot, S. 1993, ApJ, 405, 538
Bunker, A. J., Stanway, R., Ellis, R. S., & McMahon, R. G. 2004, MNRAS, 355, 374
Calzetti, D. 1997, in AIP Conf. Proc. 408, The Ultraviolet Universe at Low and High Redshift, ed. W. H. Waller et al. (New York: AIP), 403
Casertano, S., et al. 2000, AJ, 120, 2747
Cimatti, A., et al. 2002, A&A, 381, L68
Dickinson, M., et al. 2004, ApJ, 600, L99
Eales, S., Lilly, S., Webb, T., Dunne, L., Gear, W., Clements, D., & Yun, M. 2000, AJ, 120, 2244
Ferguson, H. C., Dickinson, M., & Papovich, C. 2002, ApJ, 569, L65
Ferguson, H. C., et al. 2004, ApJ, 600, L107
Gabasch, A., et al. 2004, A&A, 421, 41
Giavalisco, M. 2002, ARA&A, 40, 579
Giavalisco, M., & Dickinson, M. 2001, ApJ, 550, 177
Giavalisco, M., Steidel, C. C., Aderlberger, K. L., Dickinson, M. E., Pettini, M., & Kellogg, M. 1998, ApJ, 503, 543
Giavalisco, M., et al. 2004, ApJ, 600, L103
Gwyn, S. D. J., & Hartwick, F. D. A. 1996, ApJ, 468, L77
Hughes, D. H., et al. 1998, Nature, 394, 241
Iwata, I., Ohta, K., Tamura, N., Ando, M., Wada, S., Watanabe, C., Akiyama, M., & Aoki, K. 2003, PASJ, 55, 415
Jenkins, A., Frenk, C. S., White, S. D. M., Colberg, J. M., Cole, S., Evrard, A. E., Couchman, H. M. P., & Yoshida, N. 2001, MNRAS, 321, 372
Kennicutt, R. C., Jr. 1998, ARA&A, 36, 189
Lehnert, M. D., & Bremer, M. 2003, ApJ, 593, 630
Lilly, S. J., Le Févre, O., Hammer, F., & Crampton, D. 1996, ApJ, 460, L1
Lilly, S. J., Tresse, L., Hammer, F., Crampton, D., & Le Févre, O. 1995, ApJ, 455, 108
Lowenthal, J. D., et al. 1997, ApJ, 481, 673
Madau, P. 1995, ApJ, 441, 18
Madau, P., Ferguson, H. C., Dickinson, M. E., Giavalisco, M., Steidel, C. C., & Fruchter, A. 1996, MNRAS, 283, 1388
Madau, P., Pozetti, L., & Dickinson, M. 1998, ApJ, 498, 106
Meurer, G., Heckman, T. M., Lehnert, M., Leitherer, C., & Lowenthal, J. 1997, AJ, 114, 54
Nagamine, K., Springel, V., Hernquist, L., & Machacek, M. 2004, MNRAS, 350, 385
Oke, J. B. 1974, ApJS, 27, 21
Ouchi, M., Yamada, T., Kawai, H., & Ohta, K. 1999, ApJ, 517, L19
Ouchi, M., et al. 2004a, ApJ, 611, 660
———. 2004b, ApJ, 611, 685
Papovich, C., Dickinson, M., & Ferguson, H. C. 2001, ApJ, 559, 620
Papovich, C., et al. 2004, ApJ, 600, L111
Pettini, M., Kellogg, M., Steidel, C. C., Dickinson, M., Adelberger, K. L., & Giavalisco, M. 1998, ApJ, 508, 539
Pettini, M., Rix, S., Steidel, C. C., Adelberger, K. L., Hunt, M. P., & Shapley, A. E. 2002, ApJ, 569, 742
Pettini, M., Shapley, A., Steidel, C. C., Cuby, J.-G., Dickinson, M., Moorwood, A. F., Adelberger, K. L., & Giavalisco, M. 2001, ApJ, 554, 981
Salpeter, E. E. 1955, ApJ, 121, 161
Sawicki, M. 2002, AJ, 124, 3050
Sawicki, M., Stevenson, M., Barrientos, L. F., Gladman, B., Mallén-Ornelas, G., & van den Bergh, S. 2005, ApJ, 627, 621
Sawicki, M., & Thompson, D. 2005, ApJ, 635, 100 (KDF I)
Sawicki, M., & Yee, H. K. C. 1998, AJ, 115, 1329
Sawicki, M. J., Lin, H., & Yee, H. K. C. 1997, AJ, 113, 1
Schechter, P. 1976, ApJ, 203, 297
Schiminovich, D., et al. 2005, ApJ, 619, L47
Shapley, A. E., Steidel, C. C., Adelberger, K. L., & Pettini, M. 2001, ApJ, 562, 95
Shapley, A. E., Steidel, C. C., Pettini, M., & Adelberger, K. L. 2003, ApJ, 588, 65
Somerville, R. S., Primack, J. R., & Faber, S. M. 2001, MNRAS, 320, 504
Stanway, E. R., Bunker, A. J., & McMahon, R. G. 2003, MNRAS, 342, 439
Steidel, C. C., Adelberger, K. L., Giavalisco, M., Dickinson, M., & Pettini, M. 1999, ApJ, 519, 1
Steidel, C. C., Adelberger, K. L., Shapley, A. E., Pettini, M., Dickinson, M., & Giavalisco, M. 2003, ApJ, 592, 728

Steidel, C. C., Giavalisco, M., Pettini, M., Dickinson, M., & Adelberger, K. L. 1996, ApJ, 462, L17
Steidel, C. C., Shapley, A. E., Pettini, M., Adelberger, K. L., Erb, D. K., Reddy, N. A., & Hunt, M. P. 2004, ApJ, 604, 534
Thompson, D., et al. 1999, ApJ, 523, 100
Vijh, U. P., Witt, A. N., & Gordon, K. D. 2003, ApJ, 587, 533
Williams, R. J., et al. 1996, AJ, 112, 1335
Wyder, T. K., et al. 2005, ApJ, 619, L15
Yan, H., & Windhorst, R. A. 2004, ApJ, 600, L1