GALAXIES AT Z \sim 6: THE REST-FRAME UV LUMINOSITY FUNCTION AND LUMINOSITY DENSITY FROM 506 UDF, UDF-PS, AND GOODS I-DROPOUTS

R.J. Bouwens$^3$, G.D. Illingworth$^4$, J.P. Blakeslee$^3$, M. Franx$^3$

1 Based on observations made with the NASA/ESA Hubble Space Telescope, which is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS 5-26555. These observations are associated with programs #9803.

2 Observations have been carried out using the Very Large Telescope at the ESO Paranal Observatory under Program ID(s): LP168.A-0485.

3 Astronomy Department, University of California, Santa Cruz, CA 95064
4 Department of Physics & Astronomy, Washington State University, Pullman, WA 99164-2814 and
5 Leiden Observatory, Postbus 9513, 2300 RA Leiden, Netherlands.

Abstract

A very large sample of 506 i-dropputs (z \sim 6 galaxies) have been obtained from all the deep, wide-area HST ACS fields: HUDF, GOODS (enhanced by the extensive SNe search data – to be the v2.0 GOODS release), and UDF-Parallel ACS fields (UDF-Ps). With our selection criteria, we show that the contamination levels are \lesssim 8\% (i.e., \gtrsim 92\% or \sim 465 are at z \sim 6). This is the first comprehensive quantitative analysis of such a large sample at z \sim 6, only 0.9 Gyr from recombination, and is used to establish optimal measures of both the luminosity function and the luminosity density at z \sim 6, and their evolution relative to the z \sim 3 droput. A total of 122 i-dropputs (i_{1775} - z_{580})_{AB} > 1.3, (V_{606} - z_{580})_{AB} > 2.8 are detected in the HUDF to z_{580,AB} \sim 29.5 (the 8\% limit). Fits to the optical-infrared (ACS/NICMOS) colors of the HUDF dropout galaxies give mean UV slopes of \beta = -1.8. This is bluer than the dropout population at z \sim 3, suggesting lower dust content at z \sim 6. The sizes of the i-dropputs in our samples are small, with an \sim L^\star galaxy having r_{HI} \sim 0.8 kpc or \sim 0.14''

ABSTRACT

The rest-frame continuum UV luminosity density at z \sim 6. The best-fit Schechter parameters are M_{1350, AB} = -20.20 \pm 0.35, \alpha = -1.74 \pm 0.24, and \phi^* = 1.76^{+0.37}_{-0.73} \times 10^{-3} Mpc^{-3}. This luminosity function extends to M_{1350, AB} \sim -17.5 (0.04 L^*_{z=3}), the faintest limit yet reached for galaxies at high redshifts z > 2. We find strong evidence for evolution of the luminosity function between z \sim 6 and z \sim 3. The most likely way of accommodating this evolution is through a brightening (0.7 \pm 0.4 mag) of M^\star (at 82\% confidence) though less brightening is required if the faint–end slope \alpha changes to \alpha \sim -1.9 (from -1.6 at z \sim 3). Scenarios, such as density evolution (\phi^*), which do not include this evolution in M^\star or \alpha are ruled out at 99.99999\% confidence – demonstrating quite significantly that galaxies at z \sim 6 are lower in luminosity than galaxies at z \sim 3. Integrated down to 0.04L^*_{z=3}, the rest-frame continuum UV luminosity density (unextinguished star formation rate) at z \sim 6 is a very well-determined 1.46 \pm 0.14 \times 10^{26} ergs/s/Hz/Mpc^3. Recent determinations of the luminosity density at z \sim 6 have shown large variations – we identify some of the causes and find that other results are consistent when appropriately corrected. The substantial evolution in the luminosity function is not accompanied by a large change in the luminosity density. The luminosity density at z \sim 6 (to 0.04L^*_{z=3}) is 0.68 \pm 0.08\times that at z \sim 3. Accounting for the suggested evolution in dust content over this range indicates that the true evolution is substantially larger than for the unextinguished star formation rate, so that the total star formation rate density at z \sim 6 is just \sim 0.3\times the z \sim 3 value. Despite large uncertainties in the escape fraction and other assumed quantities, our best-fit UV luminosity function is consistent with z \sim 6 galaxies providing the necessary UV flux to reionize the universe. Lower luminosity galaxies, in particular, appear to be important for this process. Subject headings: galaxies: evolution — galaxies: high-redshift

1. INTRODUCTION

The deep z_{580}-band capabilities of the Hubble Space Telescope (HST) Advanced Camera for Surveys (ACS) have greatly enhanced the ability of astronomers to identify and observe galaxies at z \sim 6. Flux in the z_{580}-band can be contrasted with flux in the i_{1775}-band, allowing for identification of z \sim 6 i_{1775}-dropouts. Early studies revealed that i-dropputs were both smaller (Bouwens et al. 2003b; Stanway et al. 2004b; Bouwens et al. 2005c, hereinafter, B05c) and less numerous than dropouts at lower redshifts (Stanway et al. 2003; Bouwens et al. 2003b; Dickinson et al. 2004; Stanway et al. 2005; B05c). However, since much of the early work was at bright magnitudes (z_{580,AB} \lesssim 27), it was still very much unclear from these studies how this population extended to fainter magnitudes or lower surface brightnesses.

With the availability of significantly deeper i and z
data from the UDF (Beckwith et al. 2005) and the UDF parallel ACS fields (Bouwens et al. 2004b; hereinafter, UDF-Ps), this situation is largely changed. There are already a number of papers which take advantage of this depth to comment on the faint-end slope (Bouwens et al. 2004a; Bunker et al. 2004, hereinafter, BSEM; Yan & Windhorst 2004b), the rest-frame UV colors (Stanway et al. 2005; Yan et al. 2005), and the surface brightness distribution at z ~ 6 (BSEM; Bouwens et al. 2004b). These data have also provided us with some new insight into the long standing question of how the universe was reionized. Some authors (e.g., Yan & Windhorst 2004b; Lehnert & Bremer 2003) argue that it is largely the faint galaxies which were instrumental in this process, while others have emphasized the role that possible evolution in metallicity or the initial mass function (IMF) may have on the process (Stiavelli et al. 2004b). Finally, other groups (e.g., BSEM) have even questioned whether the observed galaxy population is sufficient to reionize the universe at all.

While providing many interesting early results, these early papers on i-dropouts were not comprehensive studies. Some (e.g., BSEM) restricted themselves to a bright limit (in their analyses of the two most notable data sets) to minimize the importance of incompleteness, flux, or contamination corrections. Other analyses (e.g., Yan & Windhorst 2004b) did not use a consistent i-dropout criterion across all data sets under consideration or a consistent treatment for magnitude biases, contamination, or incompleteness. Moreover, none of these early studies made an attempt to correct for possible field-to-field variations, which can be substantial (~35% RMS) for single 12 arcmin$^2$ ACS WFC fields (Somerville et al. 2004; Bouwens et al. 2004a; BSEM). Correcting for these variations is important for ensuring that a consistent normalization is used at bright, intermediate, and faint magnitudes and thus the derived LF is not compromised. Particularly important in this regard are the implications for the faint-end slope and the number of lower luminosity galaxies. Such objects have the potential to provide the necessary UV flux to reionize the universe (Lehnert & Bremer 2003; Yan & Windhorst 2004b).

The purpose of this paper is to redress many of these limitations and provide a systematic analysis of i-dropouts from some of the deepest, widest area surveys available for study. We consider fields at three different depths. The two wide-area GOODS fields (~170 arcmin$^2$ each), here enhanced to include the extensive supernova search data, form the backbone of our probe, providing important statistics at the bright end while controlling for field-to-field variations. At the faint end, there is the Hubble Ultra Deep Field (12 arcmin$^2$), which in addition to constraining the faint-end slope allows us to quantify the incompleteness, flux biases, and contamination in our shallower probes. Finally, at intermediate magnitudes, we have the two UDF-Parallel ACS fields (20 arcmin$^2$ in total), which provide an important bridge between our faintest and brightest fields. Together these three data sets provide a good measure of the i-dropout surface density over a 5 magnitude baseline, from $z_{850,AB} \sim 24.5$ to $z_{850,AB} \sim 29.5$.

This paper is structured as follows. We begin with a description of the data (§2), describe our selection criteria (§3), and then compile an i-dropout sample in the Hubble Ultra Deep Field. We use the color information to make inferences about the contamination rate, intrinsic colors, and overall redshift distribution. We then proceed to an analysis of our shallower fields and incorporate the data from those fields into our i-dropout probe, enabling an estimate of the i-dropout surface density from $z_{850,AB} \sim 24.5$ all the way down to $z_{850,AB} \sim 29.5$. In §4, we compare the present probe with previous catalogs and surface density determinations. In §5, we use this surface density to derive a LF in the rest-frame UV (~1350Å) and compare it with the LF derived at z ~ 3 (Steidel et al. 1999). Finally, we discuss these results, comment on the likely physical implications (§6), and conclude (§7). We make use of appendices to develop some key technical issues, while not interrupting the flow of the paper. Motivated by the Wilkinson Microwave Anisotropy Probe (WMAP) observations, we assume $(\Omega_M, \Omega_{\Lambda}, h) = (0.3, 0.7, 0.7)$ (Bennett et al. 2003).

2. OBSERVATIONS

As noted above, the present analysis leverages data sets of three different depths to obtain a fairly optimal measure of the number densities of i-dropouts over a 5 magnitude baseline. Table 1 provides a summary of these data sets.

2.1. ACS Ultra Deep Field

The $B_{435}V_{606}i_{775}z_{850}$ images used for this analysis are the v1.0 reductions of the UDF (Beckwith et al. 2005), binned on a 0.03′′ pixel scale. While the observations cover ~15 arcmin$^2$, our search area was restricted to the deepest 13.2 arcmin$^2$. The zeropoints used for these images are the latest values from the continuing ACS calibrations (Sirianni et al. 2005). Photometry performed using these zeropoints were offset slightly to account for the estimated Galactic absorption $E(B - V) = 0.007$ (Schlegel, Finkbeiner, & Davis 1998). The 10σ limits for these images were 29.6, 30.0, 29.9, and 29.2, respectively, in a 0.2′′-diameter aperture. PSFs were 0.09-0.10′′FWHM.

Extremely deep NICMOS coverage is available over a portion of the UDF (5.76 arcmin$^2$) (Thompson et al. 2005). That program included 8 orbits in the NIC3 $J_{110}$ filter and 8 orbits in the NIC3 $H_{160}$ filter over 9 separate pointings, for a total of 144 orbits. The pointings were arranged in a 3 × 3 grid, each separated by 45′′. Though there is some variation in depth across the mosaic, typical 5σ limits for the images were 27.6 and 27.4 in the $J_{110}$ and $H_{160}$ passbands (0.6′′-diameter aperture), respectively. Our reduction of the NICMOS data was a slight improvement on that initially made available with the treasury release and was made possible by more exact position matching with the UDF $z_{850}$-band image. This reduction is described in more detail in Thompson et al. (2005). The resulting PSFs had FWHMs of 0.33″ and 0.37″ in the $J_{110}$ and $H_{160}$ bands, respectively. The zeropoints used for the $J_{110}$ and $H_{160}$ NICMOS images were those made available by STScI (06/2004). They are 23.37 and 23.17 in $J_{110}$ and $H_{160}$, respectively. We note however that a number of workers have informally suggested that these zeropoints are too faint by ~0.2 mag. The best evidence for this comes from simultaneous fits to the ACS + NICMOS fluxes (Coe et al., in preparation). Work is ongoing at STScI to determine if there...
for this was to bin the data on a very similar 0.03
al. 2004; Bouwens et al. 2004a). Our principal reason
described in several previous publications (Blakeslee et
of these fields are used in the present paper than those
fore drizzling the images together. Different reductions
the “figure eight” patterns (resulting from scattered light
Artifacts in the original exposures like satellite trails or
formed by the “Apsis” pipeline (Blakeslee et al. 2003).

2.2. UDF ACS Parallels
The two UDF-Parallel ACS fields (UDF-Ps) were taken
in parallel to the UDF NICMOS observations (GO-9803:
Thompson et al. 2005). Each field consists of 72 orbits of
ACS observations (9 orbits $B_{435}$, 9 orbits $V_{606}$, 18 orbits
$i_{775}$, 27 orbits $z_{850}$, and 9 orbits G800L) and reach nearly
$\sim 1$ mag deeper than the original 5-epoch ACS GOODS
observations. They also reach fainter ($\sim 0.2$ − $0.4$ mags)
than the WFPC2 HDF-N (Williams et al. 1996) and
HDF-S (Williams et al. 2000). Processing of the data
included alignment, background subtraction, cosmic ray
subtraction, and drizzling onto a 0.03” grid, and was performed
by the “Apsis” pipeline (Blakeslee et al. 2003). Artifacts in the original exposures like satellite trails or
the “figure eight” patterns (resulting from scattered light
off the internal dewars) were explicitly masked out be-
before drizzling the images together. Different reductions
of these fields are used in the present paper than those
described in several previous publications (Blakeslee et
al. 2004; Bouwens et al. 2004a). Our principal reason
for this was to bin the data on a very similar 0.03”-pixel
scale to that available for the ACS GOODS fields (§2.3)
and the UDF (Beckwith et al. 2005). The similar pixel
scale made it straightforward to degrade the deeper data
down to the quality of the shallower data and therefore
estimate quantities like the completeness, flux biases, and
contamination rate (see Appendix C).

To maximize depth, we combined the ACS parallel
data (Thompson et al. 2005) with overlapping ACS WFC
exposures from the CDF-S GOODS (Giavalisco et al.
2004a), GEMS (Rix et al. 2004), and SNe search pro-
grams (Riess et al. 2005, in preparation). Incorporating
the latter data resulted in modest increases in the mean
depth of our images (+0.2 mag). Only areas having ex-
posure times in excess of 5 orbits, 11 orbits, and 18 orbits
in the $V_{606}$, $i_{775}$, and $z_{850}$ bands, respectively, are
considered in our selection (or equivalently their 10σ
depths were required to exceed 28.9, 28.6, and 28.1 in the
$V_{606}$, $i_{775}$, and $z_{850}$ bands, respectively, in 0.2”-diameter aper-
tures). This corresponded to 11.7 arcmin$^2$ in the first
UDF-Parallel [hereafter, referred to as UDFP1] and 8.2
arcmin$^2$ in the second [hereafter, referred to as UDFP2].
The 10σ depths for the deepest portion of these parallels
were 28.9, 29.2, 28.8, and 28.5 in the $B_{435}$, $V_{606}$, $i_{775}$,
and $z_{850}$ bands, respectively, in 0.2”-diameter apertures
($\sim 0.7$ − $1.1$ mags less deep than the UDF).

2.3. ACS GOODS
The current analysis makes use of our own reductions
of the ACS data available over the two GOODS areas
($\sim 170$ arcmin$^2$). Though a public reduction of the data
over this area was available (i.e., the GOODS v1.0 reduc-
tion: Giavalisco et al. 2004a), it did not include the signif-
icant amounts of ACS data taken over these fields after
the initial 398-orbit GOODS campaign. These include
195 orbits of $V_{606}$,$i_{775}$,$z_{850}$ data taken for additional SNe
searches (Riess et al. 2005, in preparation; Perlmutter et
al. 2005, in preparation), $\geq$ 100 orbits of $z_{850}$-band data
for SNe follow-up (Riess et al. 2005, in preparation; Perl-
mutter et al. 2005, in preparation), $\sim$ 40 orbits of over-
lapping $V_{606}$ and $z_{850}$ data from the GEMS program (Rix
et al. 2004), and 128 orbits of $B_{435}$,$V_{606}$,$i_{775}$,$z_{850}$ data over
the ACS parallels to the UDF NICMOS field (Thomp-
sion et al. 2004). These data substantially enhance the
GOODS v1.0 data set, and should largely be included

<table>
<thead>
<tr>
<th>Passband</th>
<th>Detection Limits (10σ)$^a$</th>
<th>PSF FWHM (&quot;)</th>
<th>Areal Coverage (arcmin$^2$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$B_{435}$</td>
<td>28.9</td>
<td>0.09</td>
<td>19.9$^b$</td>
</tr>
<tr>
<td>$V_{606}$</td>
<td>29.2</td>
<td>0.09</td>
<td>19.9$^b$</td>
</tr>
<tr>
<td>$i_{775}$</td>
<td>28.8</td>
<td>0.09</td>
<td>19.9$^b$</td>
</tr>
<tr>
<td>$z_{850}$</td>
<td>28.5</td>
<td>0.10</td>
<td>19.9$^b$</td>
</tr>
</tbody>
</table>

$^a$0.2”-diameter aperture for the ACS data, 0.6”-diameter aperture for NICMOS data, and 0.8”-diameter aperture for ISAAC data.

$^b$A significant fraction of the area from the UDF-Parallel ACS fields were not used because it did not meet our minimal S/N requirements (§2.2). The area used is tabulated here.
in the GOODS v2.0 release. Instead of waiting for this, we carried out our own reduction. Similar to our handling of the UDF-Parallel ACS fields, we have processed the ACS data with our “Apsis” pipeline (Blakeslee et al. 2003). They were drizzled onto the same astrometric grid as the images (35 individual 8k x 8k frames) which made up the v1.0 reductions of the two GOODS fields (Giavalisco et al. 2004a). These images—and our own reductions—were done on a very similar 0.03” pixel scale to the UDF. The approximate 10σ depths of those data were 28.2, 28.4, 27.7, and 27.5 in the $B_{435}$, $V_{606}$, $i_{775}$, and $z_{850}$ bands, respectively. These data reach nearly $\sim 0.15$ mag and $\sim 0.4$ mag deeper in the $i_{775}$ and $z_{850}$ bands, respectively, than the GOODS v1.0 reductions.

One complication with the analysis of the two GOODS fields is the notable variation in the depth: the extensive overlap regions between the chips (\sim 7 arcmin$^2$ for each field) being appreciably deeper ($\sim 0.4$ mag), the many outer regions ($\sim 30$ arcmin$^2$ for each field) only covered by 3 epochs of data being shallower ($\sim 0.3$ mag), and other regions of these fields with missing exposures (e.g., due to guide star acquisition problems) also being shallower (Giavalisco et al. 2004a). As we demonstrated in an earlier study on the i-dropouts from the RDCS1252-2927 (Bouwens et al. 2003b), such variations can have a dramatic impact on the number of i-dropouts selected (changing the numbers by factors of $\sim 1.8$ for just $\sim 0.4$ mag alterations in depth), and therefore any selection of i-dropouts off the degraded GOODS images requires an accurate accounting for these variations. This could be done, for example, by laying down objects at random positions across the GOODS mosaic and then attempting to recover them. Instead of adopting this more involved approach, we took a simpler route, degrading the entire frame to a uniform S/N and ignoring regions below this S/N in the object selection. Our procedure for executing the degradation is detailed in Appendix B. The threshold we settled upon was 0.1 mag brighter than that obtained with a 2.5-orbit, 3.5-orbit, and 9-orbit exposure in the $V_{606}$, $i_{775}$, and $z_{850}$ bands, respectively, and was chosen as a compromise between depth and area. This threshold is equivalent to 10σ depths of 28.3, 27.5, and 27.4 in the $V_{606}$, $i_{775}$, and $z_{850}$ bands, respectively. Throughout this work, this is what we mean when we refer to the S/N levels of the GOODS fields (or to deeper fields degraded to GOODS depth).

We also made use of the ISAAC $JK_s$ data for the CDF-South GOODS field (Vandame et al. 2005) to better estimate the contamination from lower redshift interlopers. The data consist of 21 separate \sim 3–4 hour $2.5' \times 2.5'$ ISAAC exposures in the $J$ ($\sim 1.25$μm) and $K_s$ ($\sim 2.16$μm) bands that reach $\sim 25.7$ and $\sim 25$ AB magnitudes (5σ), respectively, in a 0.8″-diameter aperture. The entire mosaic covers 131 arcmin$^2$ or about $\sim 85\%$ of the 5-epoch ACS GOODS area. Vandame et al. (2005) estimated the seeing for the frames to range from 0.31″ to 0.66″, with a median value of 0.46″. Zeropoints for the individual ISAAC frames were derived by matching photometry of \sim 50 stars on each frame with the shallower SOFI (Arnouts et al. 2001) and 2MASS (Skrutskie et al. 1997) images.

As a check on the zeropoints, we performed photometry on the $\sim 20$ $z \sim 0.4 – 1.0$ E/S0s in each ISAAC frame ($ACS\ BViz + ISAAC\ JKS$ bands) and then fit an SED to the 6 optical-infrared fluxes. While our $K_s$-band fluxes are in excellent agreement with the fit results, we noticed that our $J$-band fluxes were generally $\sim 0.1$ mag fainter than expected. Since Vandame et al. (2005) noted a similar $\sim 0.1$ mag faintward offset relative to the photometry of the K20 survey (Cimatti et al. 2002), we took this offset to be real and offset the $J$-band zeropoints quoted by Vandame et al. (2005) by 0.1 mag. No such shifts were applied to the $K$-band fluxes. Similarly, the seeing estimates obtained for different ISAAC images (Vandame et al. 2005) were examined and compared with our own estimates. In general, the FWHMs we obtained were $\sim 0.02''$ to $\sim 0.05''$ larger than the Vandame et al. (2005) estimates. We elected to apply our estimates throughout in determining the optical-infrared colors of objects in the CDF-South GOODS field (i.e., Appendix D.4.1). Given that our only use of $J$ and $K_s$ photometry in this study is for quantifying contamination, these adjustments should have no large effect on the other quantities derived here.

3. ANALYSIS

Our procedure for doing object detection and photometry is identical to that detailed in a number of previous publications by our group (e.g., Bouwens et al. 2003a; Bouwens et al. 2005b, hereinafter, B05b). SExtractor (Bertin & Arnouts 1996) was run in double-image mode, with the $z_{850}$-band images used for object detection and the other images used as the measurement images. The infrared coverage—though superior in probing beyond the break—was not used in the detection procedure because (1) the S/N and resolution of these images were in general much poorer than the $z_{850}$-band images and (2) these images—where available—tended to be very inhomogeneous in nature. Photometry was done using two scaled Kron (1980) apertures, the smaller ones to measure colors and the larger ones to convert these colors to total magnitudes. Small corrections were applied to the total magnitudes (0.1 mag to the $B_{435}$, $V_{606}$, and $i_{775}$ bands and 0.125 mag in the $z_{850}$ band: Sirianni et al. 2005) to account for the flux which falls outside these apertures (typically $\sim 0.8''$ in diameter). Optical-infrared colors were obtained by degrading the optical images to the same PSF as the coincident infrared image and then measuring the flux in an aperture which maximized the signal-to-noise (typically 0.8″-1.4″-diameter apertures).

One minor issue in the construction of our i-dropout catalogs was the choice of the SExtractor deblending parameter. A small value for this parameter will minimize blending with foreground galaxies, but will also cause many of the more clumpy i-dropouts to split into multiple pieces. Conversely, a large value for this parameter will largely avoid such splitting, but will result in more blending with foreground sources. After extensive testing, we elected to use a larger value for the deblending parameter (i.e., DEBLEND_MINCONT = 0.15) than the defaults (i.e., DEBLEND_MINCONT = 0.005). Though this will result in a greater degree of blending (e.g., 17% of i-dropouts are blended with foreground objects in the UDF vs. 11% using much smaller deblending parameters: Appendix D.1), it should avoid splitting physically-associated systems into multiple pieces—which would result in small systematic errors. Corrections can be made
for these additional incompleteness levels (see Appendix D.1).\footnote{Even better results could have been obtained here, if there was some source detection and photometry software available which had been designed to take advantage of color information in source deblending. Since dropouts have highly unique colors, it would be fairly straightforward to distinguish clumps that make up one of these objects from other foreground objects. SExtractor currently only uses the detection image for this process and does not consider color information.} To ensure that the object blending was reasonable, a detailed visual inspection was performed on each of the objects in our samples (§3.2; §3.4). No objects were found that included any obvious contribution from foreground sources. The highly unique colors of dropout sources made this check a fairly unambiguous process.

3.1. \textit{i}-dropout Selection

As in several previous publications on this subject (Stanway et al. 2003; Bouwens et al. 2004a; Dickinson et al. 2004; Bouwens et al. 2005b), \textit{i}-dropouts are selected using a simple \((i_{775} - z_{850})_{AB} < 27\) cut. At intermediate magnitudes \((24 < z_{850,AB} < 27)\), such cuts have already been shown to be quite efficient at isolating objects with blue \(z - J\) colors indicative of \(z \sim 6\) starbursts (Stanway et al. 2003; Bouwens et al. 2003b; Dickinson et al. 2004; Stanway et al. 2005; Bouwens et al. 2005b). Our choice of a more inclusive \((i_{775} - z_{850})_{AB} > 1.3\) criterion rather than the \(> 1.4\) and \(> 1.5\) criteria used in previous works (Bouwens et al. 2003; Bouwens et al. 2004a; Bouwens et al. 2005b) was motivated to maximize the size of our sample. While this will also result in a somewhat higher contamination rate, an increasing amount of data is now available, both in the IR (Thompson et al. 2005; Vandame et al. 2005) and with the ACS GRISM (Pirzkal et al. 2004; Malhotra et al. 2005) to better constrain the contamination. In addition to our \(i - z > 1.3\) criteria, we also required that objects have \((V_{606} - z_{850})_{AB}\) colors redder than 2.8 or be non-detections \((< 2\sigma)\) in the \(V_{606}\)-band to exclude lower-redshift interlopers. Appendix A provides a justification for the \((V_{606} - z_{850})_{AB}\) color cut by comparing against a number of intrinsically red galaxies uncovered in the CDF-South (Table C14). To guard against spurious sources which come in the form of low-surface brightness variations in the background (Appendix D.4.4), we required that objects in the UDF be at least 3.5\(^\prime\) detections in a 0.3\(^\prime\) aperture. The detection requirement was increased to 4 and 4.5 \(\sigma\) for the UDF-Ps and GOODS fields, respectively, to cope with the likely larger non-Gaussian signatures present in the smaller exposure stacks that comprise these data. Point sources brighter than some fiducial \(z_{850}\)-band magnitude (26.8 for GOODS fields, 27.5 for the UDF-Ps fields, and 28.4 for the UDF) were removed at this stage (point sources were defined to have SExtractor stellarity parameters \(> 0.75\)). Faintward of these fiducial limits, point sources could no longer be reliably identified (their contribution was treated as a contamination fraction and estimated statistically: see Appendix D.4.3). Table 2 contains a list of all objects excluded as stars. Finally, we carefully inspected all of our candidate \(i\)-dropouts to ensure that they did not arise from diffraction spikes around stars or the extended low-surface brightness wings around ellipticals.

3.2. \textit{i}-dropouts in the UDF

Applying the above selection criteria to the UDF results in a sample of 122 \(i\)-dropouts. Objects range in magnitude from \(z_{850,AB} = 25.0\) to \(z_{850,AB} = 29.4\) (the 8\% limit). At \(z \sim 6\), this corresponds to \(0.04 \times z\) the characteristic rest-frame UV luminosity at \(z \sim 3\) (Steidel et al. 1999). Table 3 summarizes the positions, magnitudes, \(i - z\) colors, sizes, stellarieties, \(z - J_{110}\) colors, and \(J_{110} - H_{160}\) colors of different objects in our UDF \(i\)-dropout sample. \(V_{606,i_{775},z_{850}}\) color cutouts are provided in Figure 1 for the brightest 28 \(i\)-dropouts from the UDF. Gray-scale \(z_{850}\)-band images are available for the entire catalog in Figure 2.

The deeper optical and infrared imaging available in the central region of the UDF allow us to extend our knowledge of the contamination rate of our sample from low-redshift interlopers (e.g., dusty/evolved \(z \sim 1 - 3\) objects) down to fainter magnitudes \((z_{850,AB} \geq 27)\) than has been previously possible. While there have already been several studies using these data to argue that this contamination is small (Yan & Windhorst et al. 2004b; Stanway et al. 2005), the present selection pushes slightly deeper. As in our analysis of the \(i\)-dropouts in RDCS1252-2927 and the CDF-S GOODS field (Bouwens et al. 2003b; Bouwens et al. 2005b), we consider the canonical \((i_{775} - z_{850})\) versus \((z_{850} - J_{110})\) color-color plot (Figure 3). It is immediately apparent that the contamination rate is low. Only two of the 43 \(i\)-dropouts observed to \(z_{850,AB} \sim 28.7\) had \(z - J_{110}\) colors inconsistent with the suggested position of the sample in color-color space (\textit{shaded orange region}), suggesting a very low (\(\sim 5\%\)) contamination rate for the sample as a whole. Splitting the sample across several magnitude bins, we can obtain a magnitude-dependent contamination fraction (Table D21).

3.3. Rest-frame UV Colors and Redshifts

The \(zJH\) photometry available for the UDF can also be used to estimate both the rest-frame UV colors and redshifts for sample objects. The \(z - J / J - H\) color-color diagram, in particular, serves as a useful starting point because at \(z \geq 5.9\) it provides a fairly unique mapping onto redshift and rest-frame UV color (Figure 4). Here we only include \(i\)-dropouts to a limiting magnitude of...
TABLE 2

<table>
<thead>
<tr>
<th>Object ID</th>
<th>RA</th>
<th>Dec</th>
<th>z850</th>
<th>i − z</th>
<th>z − J</th>
<th>z − K_s</th>
<th>S/G</th>
<th>r_h60''</th>
</tr>
</thead>
<tbody>
<tr>
<td>HDFN-6581218516</td>
<td>12:36:58.12</td>
<td>62:18:51.6</td>
<td>23.02±0.01</td>
<td>1.5</td>
<td>—</td>
<td>0.93</td>
<td>0.09</td>
<td></td>
</tr>
<tr>
<td>CDFS-2299552287</td>
<td>03:32:29.59</td>
<td>-27:52:28.7</td>
<td>23.05±0.01</td>
<td>1.4</td>
<td>1.0</td>
<td>0.98</td>
<td>0.08</td>
<td></td>
</tr>
<tr>
<td>HDFN-7340515534</td>
<td>12:37:34.05</td>
<td>62:15:53.4</td>
<td>23.17±0.01</td>
<td>1.5</td>
<td>—</td>
<td>0.94</td>
<td>0.09</td>
<td></td>
</tr>
<tr>
<td>CDFS-2192345455</td>
<td>03:32:19.23</td>
<td>-27:45:45.5</td>
<td>23.30±0.01</td>
<td>1.3</td>
<td>0.8</td>
<td>0.96</td>
<td>0.08</td>
<td></td>
</tr>
<tr>
<td>CDFS-2181947466</td>
<td>03:32:18.19</td>
<td>-27:47:46.6</td>
<td>23.60±0.01</td>
<td>1.4</td>
<td>1.1</td>
<td>0.99</td>
<td>0.08</td>
<td></td>
</tr>
<tr>
<td>HDFN-6388514511</td>
<td>12:36:38.85</td>
<td>62:14:51.1</td>
<td>23.89±0.01</td>
<td>1.7</td>
<td>—</td>
<td>0.92</td>
<td>0.09</td>
<td></td>
</tr>
</tbody>
</table>

aTable 2 is published in its entirety in the electronic version of the Astrophysical Journal. A portion is shown here for guidance regarding its form and content. Similar comments to Table 5 apply.

bThe same object.

TABLE 3

<table>
<thead>
<tr>
<th>Object ID</th>
<th>RA</th>
<th>Dec</th>
<th>z850</th>
<th>i − z</th>
<th>z − J</th>
<th>J − H</th>
<th>S/G</th>
<th>r_h60''</th>
</tr>
</thead>
<tbody>
<tr>
<td>UDF-40018149</td>
<td>03:32:40.01</td>
<td>-27:48:14.9</td>
<td>24.99±0.01</td>
<td>1.6</td>
<td>-0.2</td>
<td>-0.0</td>
<td>0.03</td>
<td>0.16</td>
</tr>
<tr>
<td>UDF-36476414</td>
<td>03:32:36.47</td>
<td>-27:46:41.4</td>
<td>26.08±0.02</td>
<td>2.4</td>
<td>-0.3</td>
<td>0.5</td>
<td>0.03</td>
<td>0.19</td>
</tr>
<tr>
<td>UDF-32617540</td>
<td>03:32:32.61</td>
<td>-27:47:54.0</td>
<td>26.21±0.03</td>
<td>1.4</td>
<td>—</td>
<td>—</td>
<td>0.03</td>
<td>0.24</td>
</tr>
<tr>
<td>UDF-34096472</td>
<td>03:32:34.09</td>
<td>-27:46:47.2</td>
<td>26.50±0.02</td>
<td>2.2</td>
<td>—</td>
<td>—</td>
<td>0.06</td>
<td>0.12</td>
</tr>
<tr>
<td>UDF-38286172</td>
<td>03:32:38.28</td>
<td>-27:46:17.2</td>
<td>26.56±0.03</td>
<td>2.7</td>
<td>-0.2</td>
<td>0.1</td>
<td>0.06</td>
<td>0.12</td>
</tr>
<tr>
<td>UDF-34287525</td>
<td>03:32:34.28</td>
<td>-27:47:52.5</td>
<td>26.58±0.05</td>
<td>1.5</td>
<td>0.1</td>
<td>-0.1</td>
<td>0.00</td>
<td>0.31</td>
</tr>
</tbody>
</table>

aTable 3 is published in its entirety in the electronic version of the Astrophysical Journal. A portion is shown here for guidance regarding its form and content. All magnitudes are AB. Right ascension and declination use the J2000 equinox. All limits are 2σ. “S/G” denotes SExtractor stellarity parameter, for which 0 = extended object and 1 = point source (objects with S/G > 0.75 and z850<28.4 are taken to be stars and thus not included here). “faint” indicates that an object was not detected (>2σ) in either of the passbands for a measured color and thus is not quoted. Typical errors on the i − z, z − J, and J − H colors were 0.3, 0.2-0.3, and 0.3-0.4 magnitudes, respectively.

$H_{160,AB} \sim 27.7$. Faintward of this, there are substantial errors on the J and H photometry for individual objects, and hence it is only possible to estimate the average colors for the population as a whole. We obtain these colors by stacking the i-dropouts in two differ-

Fig. 2.— z850-band postage stamps of all 122 i775-dropouts from the UDF. Otherwise the same as Figure 1.
UV colors, i.e., the data is such as to suggest moderately blue rest-frame expected position of high-redshift objects and used to estimate the contamination rate, see §.

Fig. 2.—

Our sample, such as the redshift and rest-frame UV slope.

Fig. 3.—$i_{775} - z_{850}$/$z_{850} - J_{110}$ color-color diagram showing the photometry of objects (squares) in the UDF NICMOS footprint. Objects undetected ($<2\sigma$) in the $V_{606}$ band or whose $(V_{606} - z_{850})_{AB}$ colors are redder than 2.8 (see Appendix A) are shown as black squares. Other objects in the photometric sample are displayed as magenta dots. The tracks made by starbursts of various UV spectral slopes $\beta$ are plotted here as a function of redshift to indicate the position of likely high-redshift i-dropouts (blue lines). For contrast, similar tracks have been included for a number of low-redshift templates to show the position of possible contaminants (red lines) along with the colors for early stellar types M0-T7 (Knapp et al. 2004; Geballe et al. 2002; Leggett al. 2002) (green hatched region). Though a $(i_{775} - z_{850})_{AB} > 1.3$ criterion is used to select the i-dropout sample, almost all $1 - z > 1.3$ objects have very blue $(z_{850} - J_{110})_{AB}$ colors, as expected for bona-fide $5.5 \leq z < 7$ star-forming objects (shaded orange region shows the expected position of high-redshift objects and used to estimate the contamination rate, see §3.2). This suggests that our optically selected sample has a very low contamination rate ($\lesssim 5\%$).

ent faint magnitude intervals $27.4 < z_{850,AB} < 27.9$ and $27.9 < z_{850,AB} < 28.9$. Despite concerns about possible errors in the NICMOS zeropoints (§2.1), the position of the data is such as to suggest moderately blue rest-frame UV colors, i.e., $\beta = -1.8$, and a mean redshift somewhat below 6.

Though illustrative, Figure 4 does not provide us with a very useful way of quantifying the mean properties of

Fig. 4.—$z_{850} - J_{110}/J_{110} - H_{160}$ color-color diagram showing the photometry of all 14 bright $H_{160,AB} < 27.7$ objects from the UDF with deep NICMOS coverage. Each of these objects was required to be at least a 2$\sigma$ detection in the $H_{160,AB}$-band. To make these measurements on even fainter dropouts, we included stacked photometry for objects in the magnitude bins $27.4 < z_{850,AB} < 27.9$ and $27.9 < z_{850,AB} < 28.9$ (red squares: §3.3). Model tracks are as in Figure 3. Flux errors are $1\sigma$ while the limits are $2\sigma$. The bright objects shown here are all clearly resolved and therefore non-stellar. Our $z_{850} - J_{110}$ and $J_{110} - H_{160}$ colors may be affected by a systematic 0.2 magnitude error in the $J_{110}$ and $H_{160}$ zeropoints (§2.1), shifting objects towards bluer $z_{850} - J_{110}$ colors. Nevertheless, the large fraction of objects with $H_{160} - J_{110}$ colors of $< 0.1$ (and $z_{850} - J_{110} < 0.0$) suggests a reasonably blue $\beta \lesssim -1.8$ population. Though a majority of the objects have $z_{850} - J_{110}$ colors suggestive of a pile-up at the lower redshift end of the window, i.e., $z \lesssim 6$, a few objects have $z_{850} - J_{110}$ colors consistent with being at higher redshift, i.e., $z \gtrsim 6.1$.

We can make these inferences more rigorous by performing some simulations. To make the simulations as realistic as possible, we project a UDF $B_{435}$-dropout sample (Bouwens et al. 2004b) to $z \sim 5 - 7$ and scale the
sizes of individual $B$-dropouts as $(1 + z)^{-1.1}$ (for fixed luminosity). This scaling is derived in §3.7 using the current data sets and is in good agreement with previous measurements (Bouwens et al. 2004a,b; B05c; Ferguson et al. 2004). The actual simulations are executed using our well-established cloning machinery (Bouwens et al. 1998a,b; Bouwens et al. 2003a,b; B05c) which handles the artificial redshifting and reselection of individual objects. $B$-dropouts are distributed in redshift (assuming no clustering) according to the product of their individual $1/V_{max}$ and the available cosmological volume. Here, three mean rest-frame UV slopes are assumed for the simulations: $\beta = -2.2$, $\beta = -1.8$, and $\beta = -1.4$. A one sigma scatter of 0.5 (in the UV slope $\beta$) is assumed for each. The results of the simulations are shown in the lower two panels of Figure 5 (black lines) and compared with the observations. It seems clear that the $\beta = -1.8$ model (dotted black line) provides the best representation of the observed $(J_{110} - H_{160})_{AB}$ colors (histogram). All models however yield a tail toward large $(z_{580} - J_{110})_{AB}$ colors which does not occur in the observations (histogram). This suggests a deficit of i-dropouts at the higher redshift end of the $z \sim 5.5 - 7$ selection window. To model this, we assumed the space density of i-dropouts in the UDF was a strong function of redshift, i.e., $e^{-\left(z-5.5\right)}$, while adopting the best-fitting mean-frame UV slope $\beta$ found above ($-1.8$). The results are shown in Figure 5 as a solid purple line and provide a rough fit to the median colors. We note that very similar conclusions have come out of the GRAPES program (Malhotra et al. 2005), where even better redshift measurements are possible from the GRISM data. Malhotra et al. (2005) demonstrate that the majority of bright $(z_{580,AB} \lesssim 28)$ i-dropouts in the UDF (15 out of 23 objects) are at $z \sim 5.9 \pm 0.2$.

3.4. i-dropouts in the GOODS/UDF-Ps fields

To control for field-to-field variations and to add numbers at bright and intermediate magnitudes (where statistics in the UDF are poor), it was useful to incorporate the UDF results with those derived from the shallower UDF-Ps and GOODS fields. The selection of dropouts from these fields was performed using nearly identical selection criteria to that used for the UDF (§3.2). 68 and 332 dropouts were found in the UDF-Ps and GOODS fields, respectively (Tables 4-5). This is significantly more dropouts ($\sim 2 - 5 \times$) than were found in our initial studies on these fields (Bouwens et al. 2004a; B05b) and owes to our slightly more inclusive selection criteria ($i - z > 1.3$ rather than $i - z > 1.4$), better pixelization (0.03″ rather than 0.05″), greater depths (0.2 mag fainter for the UDF-Ps and 0.4 mag fainter for the GOODS fields), and larger areas probed (an additional $\sim 70$ arcmin$^2$ for the GOODS fields). Our total i-dropout sample (from all three data sets) has 506 individual objects (16 of the total 522 dropouts from these three fields are found in both our GOODS and UDF/UDF-Ps catalogs and so only need to be counted once).

3.5. Corrections for Depth

The properties of all our i-dropout samples are summarized in Table 6. To put these samples together to obtain a single measure of the i-dropout surface density, we must account for the sizeable effect of survey depth. A simple illustration of this can be found in the top panel of Figure 6, which contrasts i-dropouts selected from the UDF, UDF-Ps, and GOODS fields. Though incompleteness is clearly the dominant effect in the observed differences, other selection and measurement biases also play a role. We relegate a detailed discussion of these biases to Appendix D. However, it is useful to give a brief summary here of the main corrections.

We divide these corrections into incompleteness, flux, and contamination corrections. These corrections allow an approximate conversion from the surface densities measured in our shallower data to their equivalent surface densities if measured with UDF quality data. Our first correction, the completeness corrections (Appendix D.1), makes up for the fact that our shallower surveys preferentially miss the larger, lower surface brightness fraction of galaxies in any given magnitude interval. In general, these corrections tend to be small ($\lesssim 10\%$) except near the magnitude limit of the data, where they can be $\geq 50\%$. For the GOODS data, this comes at $z_{580,AB} \gtrsim 26.8$ and for the UDF-Ps data, this comes at...
9

TABLE 4
UDF-Ps i-dropout sample.

<table>
<thead>
<tr>
<th>Object ID</th>
<th>RA</th>
<th>Dec</th>
<th>$z_{850}$</th>
<th>$i - z$</th>
<th>$z - J$</th>
<th>$z - K_s$</th>
<th>S/G</th>
<th>$r_{mj}$ (&quot;)</th>
</tr>
</thead>
<tbody>
<tr>
<td>UDFP1-2494954244</td>
<td>03:32:49.49</td>
<td>-27:54:24.4</td>
<td>26.68±0.06</td>
<td>1.4</td>
<td>0.00</td>
<td>0.21</td>
<td></td>
<td></td>
</tr>
<tr>
<td>UDFP2-2064148469</td>
<td>03:32:06.41</td>
<td>-27:48:46.9</td>
<td>26.11±0.06</td>
<td>1.7</td>
<td>0.01</td>
<td>0.19</td>
<td></td>
<td></td>
</tr>
<tr>
<td>UDFP1-2493956440</td>
<td>03:32:43.98</td>
<td>-27:56:44.0</td>
<td>26.32±0.04</td>
<td>1.6</td>
<td>0.38</td>
<td>0.10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>UDFP1-2483955541</td>
<td>03:32:48.39</td>
<td>-27:55:54.1</td>
<td>26.63±0.05</td>
<td>1.6</td>
<td>0.35</td>
<td>0.09</td>
<td></td>
<td></td>
</tr>
<tr>
<td>UDFP1-2427156555</td>
<td>03:32:42.71</td>
<td>-27:56:55.5</td>
<td>26.77±0.09</td>
<td>2.0</td>
<td>0.00</td>
<td>0.17</td>
<td></td>
<td></td>
</tr>
<tr>
<td>UDFP1-2394754149</td>
<td>03:32:39.47</td>
<td>-27:54:14.9</td>
<td>26.88±0.10</td>
<td>1.3</td>
<td>0.01</td>
<td>0.15</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Table 4 is published in its entirety in the electronic version of the Astrophysical Journal. A portion is shown here for guidance regarding its form and content. All magnitudes are AB. Right ascension and declination use the J2000 equinox. All limits are 2σ.*

TABLE 5
GOODS i-dropout sample.*

<table>
<thead>
<tr>
<th>Object ID</th>
<th>RA</th>
<th>Dec</th>
<th>$z_{850}$</th>
<th>$i - z$</th>
<th>$z - J$</th>
<th>$z - K_s$</th>
<th>S/G</th>
<th>$r_{mj}$ (&quot;)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CDFS-2256155487</td>
<td>03:32:25.61</td>
<td>-27:55:48.7</td>
<td>24.51±0.02</td>
<td>1.6</td>
<td>-0.1</td>
<td>-1.0</td>
<td>0.37</td>
<td>0.11</td>
</tr>
<tr>
<td>HDFN-5426312901</td>
<td>12:35:42.63</td>
<td>62:12:09.1</td>
<td>25.15±0.06</td>
<td>1.5</td>
<td>—</td>
<td>—</td>
<td>0.02</td>
<td>0.23</td>
</tr>
<tr>
<td>CDFS-2400148141</td>
<td>03:32:40.01</td>
<td>-27:48:14.1</td>
<td>25.17±0.04</td>
<td>1.6</td>
<td>0.0</td>
<td>0.0</td>
<td>0.36</td>
<td>0.12</td>
</tr>
<tr>
<td>CDFS-2331393491</td>
<td>03:32:33.19</td>
<td>-27:39:49.1</td>
<td>25.29±0.06</td>
<td>2.4</td>
<td>0.02</td>
<td>0.21</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CDFS-2337840378</td>
<td>03:32:23.78</td>
<td>-27:40:37.8</td>
<td>25.34±0.07</td>
<td>1.6</td>
<td>—</td>
<td>—</td>
<td>0.01</td>
<td>0.22</td>
</tr>
<tr>
<td>CDFS-2343852466*</td>
<td>03:32:33.48</td>
<td>-27:52:46.6</td>
<td>25.37±0.08</td>
<td>1.4</td>
<td>1.2</td>
<td>2.5</td>
<td>0.00</td>
<td>0.29</td>
</tr>
</tbody>
</table>

*Table 5 is published in its entirety in the electronic version of the Astrophysical Journal. A portion is shown here for guidance regarding its form and content. All magnitudes are AB. Right ascension and declination use the J2000 equinox. All limits are 2σ.*

TABLE 6
Summary of i-dropout samples.*

<table>
<thead>
<tr>
<th>Sample</th>
<th>Area</th>
<th>Mag.</th>
<th>Limit</th>
<th>$L_{z-3}^+$</th>
</tr>
</thead>
<tbody>
<tr>
<td>CDFS GOODS</td>
<td>193</td>
<td>181</td>
<td>z ~ 27.9</td>
<td>0.17</td>
</tr>
<tr>
<td>HDFN GOODS</td>
<td>175</td>
<td>151</td>
<td>z ~ 27.9</td>
<td>0.17</td>
</tr>
<tr>
<td>UDFP1</td>
<td>12</td>
<td>54†</td>
<td>z ~ 28.6</td>
<td>0.09</td>
</tr>
<tr>
<td>UDFP2</td>
<td>8</td>
<td>14†</td>
<td>z ~ 28.6</td>
<td>0.09</td>
</tr>
<tr>
<td>UDF</td>
<td>13</td>
<td>122†</td>
<td>z ~ 29.4</td>
<td>0.04</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Sample</th>
<th>Area</th>
<th>Mag.</th>
<th>Limit</th>
<th>$L_{z-3}^+$</th>
</tr>
</thead>
<tbody>
<tr>
<td>CDFS GOODS</td>
<td>193</td>
<td>181</td>
<td>z ~ 27.9</td>
<td>0.17</td>
</tr>
<tr>
<td>HDFN GOODS</td>
<td>175</td>
<td>151</td>
<td>z ~ 27.9</td>
<td>0.17</td>
</tr>
<tr>
<td>UDFP1</td>
<td>12</td>
<td>54†</td>
<td>z ~ 28.6</td>
<td>0.09</td>
</tr>
<tr>
<td>UDFP2</td>
<td>8</td>
<td>14†</td>
<td>z ~ 28.6</td>
<td>0.09</td>
</tr>
<tr>
<td>UDF</td>
<td>13</td>
<td>122†</td>
<td>z ~ 29.4</td>
<td>0.04</td>
</tr>
</tbody>
</table>

*Due to our inclusion of the ACS parallels to the UDF NICMOS field in our reductions of the CDF South GOODS field (§2.3), the total area available there for i-dropout searches exceeded that available in the HDF-North GOODS field.

The effective normalization of the luminosity function is expected to show significant variations as a function of field-to-field variations. The recent findings from the GRAPES team (Malhotra et al. 2005) are consistent with these contamination estimates. For an $z_{500, AB}$ dropout sample.

$z_{850, AB} > 27.5$. Table D17 show the results of the simulations. As with other results in this section, these results were obtained by degrading the deeper data down to the depths of the shallower data and repeating the selection. The purpose of our second set of corrections, the flux corrections (Appendix D.2), was to compensate for the fact that our shallower surveys may estimate lower fluxes for objects than would be measured in deeper exposures. Here, the corrections proved to be fairly small ($\sim 0.1–0.2$ mags) except for those objects near the magnitude limit (Figure D20) where some brightening was observed (0.3 mag). This brightening appeared to be the result of a Malmquist bias. The first and second corrections were implemented using the estimated transfer functions (Appendix D.3: Tables D19 and D20). Finally, our third set of corrections (Appendix D.4) were used to subtract out the likely contamination rate for our different samples. We included a variety of different sources of contamination in this estimate: low-mass stars (Figure D23), intrinsically-red lower redshift interlopers (Table D21), objects which entered our sample due to photometric scatter (Tables D22-D23), and finally spurious objects. In general, all sources of contamination were small and never contributed more than 15% of the objects in any given magnitude interval.2

3.6. Field-to-Field Variations

The effective normalization of the luminosity function is expected to show significant variations as a function of field-to-field variations. The recent findings from the GRAPES team (Malhotra et al. 2005) are consistent with these contamination estimates. For an $z_{500, AB}$ dropout sample.

$z_{500, AB} > 1.3$ selection (where the spectra could be unambiguously extracted), the GRAPES team found that only one out of 15 objects was a contaminant (a $z_{500, AB} \sim 25.4$ star). In the current UDF selection (§3.1), this object was rejected as a star.

Table D17 show the results of the simulations. As with other results in this section, these results were obtained by degrading the deeper data down to the depths of the shallower data and repeating the selection. The purpose of our second set of corrections, the flux corrections (Appendix D.2), was to compensate for the fact that our shallower surveys may estimate lower fluxes for objects than would be measured in deeper exposures. Here, the corrections proved to be fairly small ($\sim 0.1–0.2$ mags) except for those objects near the magnitude limit (Figure D20) where some brightening was observed (0.3 mag). This brightening appeared to be the result of a Malmquist bias. The first and second corrections were implemented using the estimated transfer functions (Appendix D.3: Tables D19 and D20). Finally, our third set of corrections (Appendix D.4) were used to subtract out the likely contamination rate for our different samples. We included a variety of different sources of contamination in this estimate: low-mass stars (Figure D23), intrinsically-red lower redshift interlopers (Table D21), objects which entered our sample due to photometric scatter (Tables D22-D23), and finally spurious objects. In general, all sources of contamination were small and never contributed more than 15% of the objects in any given magnitude interval.2

2 The recent findings from the GRAPES team (Malhotra et al. 2005) are consistent with these contamination estimates. For an $z_{500, AB} > 1.3$ selection (where the spectra could be unambiguously extracted), the GRAPES team found that only one out of 15 objects was a contaminant (a $z_{500, AB} \sim 25.4$ star). In the current UDF selection (§3.1), this object was rejected as a star.
sis representative of the cosmic average, our challenge is to derive a luminosity function which (loosely referred to as “cosmic variance”). Since the goal is to remove these variations. A simple averaging of the $i$-dropouts from the different fields is not appropriate since the fields differ in the magnitude ranges they probe. One would have no guarantee with such a procedure that the average normalization obtained at brighter magnitudes is the same as that obtained at fainter magnitudes, thus allowing for discontinuities in the normalization. This could impact the shape of the derived LF (see Appendix E).

To remove these differences, it is necessary to estimate the relative normalizations of $i$-dropouts in our different survey fields. We do this by degrading the deeper data down to the depths of the shallower survey fields and then comparing the surface densities of $i$-dropouts derived. To maximize the significance, the present comparisons are done in two stages: first comparing the UDF against the UDF-Ps and second comparing the deeper three fields (UDF + UDF-Ps) against the GOODS fields. The overall normalization for our data sets is set by the mean of the two GOODS fields, which sampling the largest comoving volume, should provide our best estimate of the cosmic average.

For the first stage, $i$-dropouts in the UDF are normalized relative to $i$-dropouts in our deepest three fields. The normalization factor is determined by degrading the UDF down to the same S/N level as the two parallels and then comparing the number of dropouts in the fields. This degradation was performed 10 times and the signal-to-noise (weight maps) of both parallels were matched on a pixel-by-pixel basis (as in Appendix C and Appendix D.1). Our findings are shown in Table 7 and point to the UDF having a similar normalization to the first parallel (50.2 vs. 43.6), but substantially higher than for the second parallel (27.8 vs. 11.4). Taken together, this suggests that the UDF is $16\pm24\%$ overdense relative to the mean of the UDF-Ps, or $10\pm15\%$ overdense relative to the cosmic average defined by the three fields. These fields also enable us to remark on the observed field-to-field variations, which appear to be $\sim46\%$ RMS. This is consistent with the $\sim35\%$ RMS variations one obtains assuming a $\Lambda$CDM power spectrum, $\Delta z = 0.7$ selection window, pencil beam geometry, and bias of 4, which is appropriate (Mo & White 1996) for objects of number density $\sim10^{-3}$ Mpc$^{-3}$ probed by these fields (Figure 11).

For the second stage, the normalization of the deeper three fields are adjusted to match that of the GOODS fields. As before, the normalization factor is estimated by degrading the UDF and UDF-Ps to the signal-to-noise level of the GOODS fields and extracting $i$-dropout samples using selection criteria identical to that used for GOODS. The results of these experiments are shown in Table 8, and it is clear that the average surface density derived from the three deeper fields ($0.59 \pm 0.13$ arcmin$^{-2}$) is somewhat lower ($0.69 \pm 0.16\times$) than $3\sigma$ Poissonian uncertainties. The top panel presents the uncorrected surface densities. The middle panel presents the same surface densities corrected for completeness, flux biases, and contamination rates (Appendix D). The bottom panel shows the cumulative surface density obtained combining these fields through a maximum likelihood procedure ($\S3.8$). The predicted surface densities of $i$-dropouts assuming the Steidel et al. (1999) LF, the Steidel et al. (1999) LF divided by 3, and the Steidel et al. (1999) LF divided by 6 are shown with the three solid black curves in the bottom panel. The uncorrected UDF counts are also included in the bottom panel as the dotted red histogram. The equivalent differential counts of BCSM are included in this panel as blue circles (from their Figure 10). The no-evolution ($z \sim 3$) predictions exceed the observed counts by factors of $\approx 6\times$ at the bright end ($z_{505, AB} \lesssim 26$: Stanway et al. 2003, 2004b; Dickinson et al. 2004), $\approx 3\times$ at more intermediate magnitudes ($z_{505, AB} \sim 26-27$), and $\lesssim 2\times$ at faint end ($z_{505, AB} \gtrsim 27$). The effect of depth on the extracted counts is obvious in the top panel. A detailed quantification of the relevant biases (selection and measurement) is provided in Table D17 and Figure D20.

Note that more objects are found in the degradation of the UDF to the depth and area of the first parallel than the second. This is due to the slightly larger depth and area for that field (due to a greater overlap with exposures from the GOODS fields).

For example, the first stage normalization factor (1.10$^{+0.15}_{-0.10}$) quoted for the UDF can be calculated from the numbers given in Table 7 as $3(35.7)/(11.4 + 50.2 + 35.7) \sim 1.10$ where 35.7 is the average number of dropouts found in the degradations of the UDF to the depth of the parallels, i.e., $\frac{3 \times 35.7}{3} \sim 35.7$.

Fig. 6.—Surface Densities (per 0.5 mag interval) of $i$-dropouts observed at three different depths: GOODS (black histogram), UDF-Ps (blue histogram), and UDF (red histogram). Errors are $1\sigma$ Poissonian uncertainties. The top panel presents the uncorrected surface densities. The middle panel presents the same surface densities corrected for completeness, flux biases, and contamination rates (Appendix D). The bottom panel shows the cumulative surface density obtained combining these fields through a maximum likelihood procedure ($\S3.8$). The predicted surface densities of $i$-dropouts assuming the Steidel et al. (1999) LF, the Steidel et al. (1999) LF divided by 3, and the Steidel et al. (1999) LF divided by 6 are shown with the three solid black curves in the bottom panel. The uncorrected UDF counts are also included in the bottom panel as the dotted red histogram. The equivalent differential counts of BCSM are included in this panel as blue circles (from their Figure 10). The no-evolution ($z \sim 3$) predictions exceed the observed counts by factors of $\approx 6\times$ at the bright end ($z_{505, AB} \lesssim 26$: Stanway et al. 2003, 2004b; Dickinson et al. 2004), $\approx 3\times$ at more intermediate magnitudes ($z_{505, AB} \sim 26-27$), and $\lesssim 2\times$ at faint end ($z_{505, AB} \gtrsim 27$). The effect of depth on the extracted counts is obvious in the top panel. A detailed quantification of the relevant biases (selection and measurement) is provided in Table D17 and Figure D20.

position and environment ($\sim35\%$ RMS for single ACS pointings). This is the result of large-scale structure (loosely referred to as “cosmic variance”). Since the goal of these studies is to derive a luminosity function which is representative of the cosmic average, our challenge is
than our estimates made with the two stage procedure. These values are compiled under the column “One Stage” in Table 9. The UDF and UDF-Ps fields, respectively. These factors are summarized in Table 8. 

<table>
<thead>
<tr>
<th>Field</th>
<th># of dropouts</th>
</tr>
</thead>
<tbody>
<tr>
<td>UDFP1</td>
<td>50.2†</td>
</tr>
<tr>
<td>UDFP2</td>
<td>11.4‡</td>
</tr>
<tr>
<td>UDF</td>
<td>43.6</td>
</tr>
</tbody>
</table>

†These numbers have been corrected for the expected contamination from low-redshift objects scattering into our sample (~3 per field, see Table D22).
‡Note that no comparable deficit in B or V-dropouts is found in UDFP2 relative to other fields (e.g., the UDF or UDFP1), suggesting the apparent underabundance of i-dropouts here is not related to the reduction or processing of the data (or any bright stars in the foreground).

The depth and selection area in the second parallel were smaller than that of the first due to a lesser overlap with GOODS. As a result, degradations of the UDF to the depth of the first parallel revealed more objects than a comparable degradation to the depth of the second.

<table>
<thead>
<tr>
<th>Field</th>
<th>Surface Density*</th>
</tr>
</thead>
<tbody>
<tr>
<td>HDFN GOODS</td>
<td>0.81±0.07†</td>
</tr>
<tr>
<td>CDF-S GOODS</td>
<td>0.89±0.07†</td>
</tr>
<tr>
<td>UDFP1</td>
<td>0.67±0.24</td>
</tr>
<tr>
<td>UDFP2</td>
<td>0.28±0.18</td>
</tr>
<tr>
<td>UDF</td>
<td>0.82±0.25</td>
</tr>
</tbody>
</table>

*Units are arcmin⁻².
†These surface densities have been corrected for the expected contamination rate from low-redshift objects scattering into our sample (0.05 contaminants arcmin⁻², see Tables D22 and D23).

A rather unexpected result is that found in both GOODS fields (0.85±0.05 arcmin⁻²). 0.69±0.16 is the second stage normalization factor. Interestingly enough, the surface density of i-dropouts is 9% ± 13% larger in the CDF-S GOODS field than it is in the HDF-N GOODS field. However, this is not inconsistent with the sort of variations expected from cosmic variance (±20%) in fields of this size (184 arcmin²) (Somerville et al. 2004). Multiplying the first and second stage factors together, we arrive at an overall normalization factor for the UDF and UDF-Ps. These factors are summarized in Table 9 under the column “Two Stage.”

As an alternative to this procedure, the normalization of our deeper fields can be derived by comparing directly with the surface density of i-dropouts found at GOODS depth (Table 8). Using the above results (i.e., Table 8), we derive a normalization of 0.96±0.30 and 0.56±0.18 for the UDF and UDF-Ps fields, respectively. These values are compiled under the column “One Stage” in Table 9. While consistent, they are of slightly lower significance than our estimates made with the two stage procedure. We adopt the results of the two stage procedure as our final estimate of the relative normalization and take the reciprocal of this normalization as our adjustment factor.

### Table 7

<table>
<thead>
<tr>
<th>Field</th>
<th>UDFP1</th>
<th>UDFP2</th>
</tr>
</thead>
<tbody>
<tr>
<td>UDFP1</td>
<td>50.2†</td>
<td>–</td>
</tr>
<tr>
<td>UDFP2</td>
<td>11.4‡</td>
<td>43.6</td>
</tr>
<tr>
<td>UDF</td>
<td>27.8*</td>
<td>–</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Relative Normalization</th>
<th>Two Stage*</th>
<th>One Stage†</th>
<th>Adj. Factor‡</th>
</tr>
</thead>
<tbody>
<tr>
<td>UDFPs</td>
<td>0.66 ± 0.16</td>
<td>0.56 ± 0.18</td>
<td>1.52</td>
</tr>
<tr>
<td>UDF</td>
<td>0.76 ± 0.20</td>
<td>0.96 ± 0.30</td>
<td>1.32</td>
</tr>
<tr>
<td>GOODS</td>
<td>1.0 (fixed)</td>
<td>1.0 (fixed)</td>
<td>1.00</td>
</tr>
</tbody>
</table>

*The two stage normalization (§3.6) is obtained by comparing the surface densities of i-dropouts in a field with those of the two GOODS fields. This is a two stage process, where the normalization of a given field is first tied to the deepest three fields (Table 7) and these fields, in turn, are tied to the two GOODS fields (Table 8). The final normalization factor is then the product of the normalization factors derived from these two comparisons, e.g., (1.10 ± 0.15)/(0.69 ± 0.16) = 0.76 ± 0.20 for the UDF (see §3.6). The two stage normalization has the advantage of a larger overlap between the different surveys being linked up. This overlap translates into smaller uncertainties in the overall normalization factors (estimated assuming poissonian errors).
†The one stage normalization (§3.6) is obtained by comparing the surface densities of i-dropouts in a field with the average of that found in the two GOODS fields (Table 8), e.g., (0.82 ± 0.25 arcmin⁻²)/(0.81 ± 0.07 arcmin⁻²) = 0.96 ± 0.30 for the UDF.
‡The adopted adjustment factor is equal to the reciprocal of the normalization relative to GOODS. We use the two stage normalizations because of their smaller uncertainties.

### 3.7. Dependence of Galaxy Size on Redshift

Data from the UDF, UDF-Ps, and GOODS fields also allow us to revisit our analyses on the physical sizes of galaxies at $z \sim 6$ and how these sizes compare with those at later times. Previously, we had carried out our analyses using each of the above fields separately (Bouwens et al. 2004b; Bouwens et al. 2004a; B05c). With the combined data set, we can significantly improve this analysis. For the present paper, these sizes are important for modelling the selection effects of our i-dropout samples. Similar to our previous works, we model the sizes of i-dropouts in all three samples using different size scalings $(1 + z)^{-m}$ of a $z \sim 2$ HDF-N + HDF-S U-dropout sample (B05b). We project objects from this sample to higher redshift ($z \sim 5 - 7$) using our cloning machinery, add them to noise frames, and then reselect them in exactly the same way as for the observed samples. Only galaxies one magnitude brighter than $i$-dropouts from all three fields. Here it is evident that the typical half-light radius for i-dropouts at $z_{50,AB} \sim 27$ is 0.8 kpc (after correction for the PSF). Relative to the sizes of objects at lower redshift, the $(1 + z)^{0}$ and $(1 + z)^{-2}$ scalings seem to nicely bracket the observed range. To derive a more precise estimate, we rely on comparisons between the mean half-light radii obtained from the observations and simulations. Interpolating between our simulation results, our best-fit values for the scaling exponent $m$ are 1.2±0.4, 1.0±0.5, and 1.0±0.4 for the UDF, UDF-Ps, and GOODS fields, respectively. Combining the results from the three fields to obtain a single scaling (and thus assuming that this redshift scaling is luminosity independent)
TABLE 10
i-dropouts surface densities estimated from UDF, UDF-Ps, and GOODS fields, corrected up to the UDF completeness levels.\textsuperscript{a}

<table>
<thead>
<tr>
<th>Surface Density (arcmin(^{-2}))</th>
<th>I</th>
<th>II</th>
</tr>
</thead>
<tbody>
<tr>
<td>24.50 &lt; (z_{850}) &lt; 25.00</td>
<td>0.003 ± 0.003</td>
<td>0.004 ± 0.004</td>
</tr>
<tr>
<td>25.00 &lt; (z_{850}) &lt; 25.50</td>
<td>0.020 ± 0.010</td>
<td>0.024 ± 0.012</td>
</tr>
<tr>
<td>25.50 &lt; (z_{850}) &lt; 26.00</td>
<td>0.047 ± 0.016</td>
<td>0.056 ± 0.019</td>
</tr>
<tr>
<td>26.00 &lt; (z_{850}) &lt; 26.50</td>
<td>0.143 ± 0.028</td>
<td>0.172 ± 0.034</td>
</tr>
<tr>
<td>26.50 &lt; (z_{850}) &lt; 27.00</td>
<td>0.473 ± 0.061</td>
<td>0.570 ± 0.073</td>
</tr>
<tr>
<td>27.00 &lt; (z_{850}) &lt; 27.50</td>
<td>1.008 ± 0.228</td>
<td>1.215 ± 0.275</td>
</tr>
<tr>
<td>27.50 &lt; (z_{850}) &lt; 28.00</td>
<td>2.222 ± 0.410</td>
<td>2.677 ± 0.494</td>
</tr>
<tr>
<td>28.00 &lt; (z_{850}) &lt; 28.50</td>
<td>4.140 ± 0.391</td>
<td>1.699 ± 0.471</td>
</tr>
<tr>
<td>28.50 &lt; (z_{850}) &lt; 29.00</td>
<td>4.061 ± 0.684</td>
<td>4.893 ± 0.824</td>
</tr>
<tr>
<td>29.00 &lt; (z_{850}) &lt; 29.50</td>
<td>1.496 ± 0.415</td>
<td>1.805 ± 0.498</td>
</tr>
</tbody>
</table>

\textsuperscript{a}Because of the modest (~17\%) incompleteness due to object blending in the HUDF (Appendix D.1), we quote two different surface densities here. Column “I” gives the equivalent surface densities at UDF depths. Column “II” corrects the column “I” surface densities for blending (i.e., by multiplying column “I” by 1/0.83). The results in column “II” should be largely free of selection or measurement biases brightward of \(z_{850,AB} \sim 28.5\). Faintward of this, incompleteness becomes important.

yields \(m = 1.1±0.3\). This is in good agreement with several previous determinations: \(m = [0.8, 2.0]_\sigma\) (B05c), \(m = 1.57±0.53\) (Bouwens et al. 2004a), \(m = 0.94_{-0.19}^{+0.25}\) (Bouwens et al. 2004b), and the Ferguson et al. (2004) \(H(z)^{-1}\) size scaling, which is equivalent to \(m = 1.47\) over the redshift range 2.5 < \(z\) < 6.

3.8. Best-Fit Surface Densities

It is useful to combine the results from our three data sets into a single measure of the i-dropout surface density as a function of magnitude. To derive this, we apply a maximum likelihood procedure. For all three data sets, the model counts are convolved with the transfer functions (Appendix D.3: from the UDF to the relevant field), multiplied by the normalization factors from Table 9 (from the cosmic average to the normalization of the particular field), and then compared with the observed counts. In these fits, we do not include counts faintward of \(z_{850,AB} = 27.0\) in the two GOODS fields and faintward of \(z_{850,AB} = 28.0\) in the UDF-Ps to be conservative. This allows us to avoid any systematics which may occur in modelling the selection effects near the completeness limit. The resultant surface density of i-dropouts is tabulated in column I of Table 10 and shown in the bottom panel of Figure 6. This surface density spans 5 magnitudes, running all the way from \(z_{850,AB} \sim 24.5\) to \(z_{850,AB} \sim 29.5\). We remind the reader that the surface densities quoted here are as measured at UDF depths and are not free of the incompleteness/flux biases implicit at these levels. Because of this, we have also included a second column in Table 10 which quotes the surface densities at UDF depths corrected for blending with foreground objects (see Appendix D.1).

4. COMPARISON AGAINST PREVIOUS RESULTS

4.1. Source Lists / Surface Densities

In the previous section, we used i-dropouts measured at three different depths (GOODS, UDF-Ps, UDF) to derive an optimal measure of the surface density of i-dropouts. Previously, there have been several attempts at compiling the counts from these fields, and so it is useful to make comparisons against the source lists first before trying to understand possible differences in the interpretation. We begin with the i-dropouts from the UDF, where several source lists have already been compiled (BSEM; Yan & Windhorst 2004b; Beckwith et al. 2005). Fortunately, these papers use nearly identical selection criteria to the present sample, facilitating the comparisons. As far as the current catalogs are concerned, 48 of the 54 i-dropouts compiled by BSEM appear in our primary list (Table 3), 4 appear in our blended \(z \sim 6\) candidate list (Table D18; see Appendix D.1), one (BSEM#49117) was blended with a foreground object in both our catalogs (Tables 3 and D18), and one (BSEM#17487) had a \(V_{606}\)-band flux \(V_{606} - z_{850} = 2.4\) inconsistent with our i-dropout selection criteria. 84 of the brightest 95 i-dropouts \((z_{850,AB} < 29.5)\) from the Yan & Windhorst (2004b) catalog also appear in our primary list (Table 3), 5 appear in our blended \(z \sim 6\) candidate list (Table D18), 3 had \(V_{606}\)-band fluxes inconsistent with our i-dropout criteria, and 3 were near the edges of the UDF image and therefore outside our selection area. Possible differences in object splitting between catalogs are ignored in the above comparisons.

As for the previously published catalogs, 35 of the brightest 39 i-dropouts from Table 3 \((z_{850,AB} < 27.9)\) appear in the BSEM catalog and 34 of these 39 appear in the Yan & Windhorst (2004b) compilation. Objects appear to be missing from the previous catalogs due to their surface brightness (e.g., as with UDF-4256665666 or UDF-34998369), proximity to the edges of the UDF image and therefore outside our selection area. Possible differences in object splitting between catalogs are ignored in the above comparisons.

In the GOODS fields, the surface densities we derive are less than those first reported by Giavalisco et al. (2004b) and Dickinson et al. (2004) using a similar \((i_{775} - z_{850})_{AB} > 1.3\) selection on the 3-epoch data. We obtain 0.08±0.02 arcmin\(^{-2}\) and 0.26±0.04 arcmin\(^{-2}\) to \(z_{850,AB} \sim 26\) and \(z_{850,AB} \sim 26.5\), respectively, versus their surface densities 0.17 arcmin\(^{-2}\) and 0.37 arcmin\(^{-2}\) to the same magnitude limits, after applying their estimated correction for contamination from photometric scatter (20%) and spurious fraction (23%). The disagreement becomes even worse, however, if an account is made for the fact that their surface densities derive from the 3-epoch data (and would need to be corrected upwards to account for the considerable incompletenesses at these depths). What is the source of this disagreement? A quick investigation suggests that it has come from a substantial underestimate of the contamination rate in these previous studies. Here we can revisit these estimates using the deeper GOODS data and especially the UDF-Ps and UDF data. Of the 251 i-dropouts in the Dickinson et al. (2004) i-dropout catalog, only 12 overlap with the deeper UDF (2 mag fainter) and UDF-Ps (1 mag fainter) data. Three (25\%) of these objects appear to be bona-fide i-dropouts, 2 (17\%) are low-redshift interlopers, and 7 objects (58\%) are not found at all in the
 deeper data and therefore appear to be spurious. This works out to a 75% contamination rate, which is much higher than the ~45% estimated in the Giavalisco et al. (2004b) and Dickinson et al. (2004) studies. To be fair, we note that these studies stressed the substantial uncertainties in their estimates. Even more striking is the fact that only 94 of the 251 i-dropouts in the Dickinson et al. (2004) catalog are found in the current catalogs (based on data which are ~0.6 mag deeper in the $z_{850}$ band) and 48 of these qualify as i-dropouts using our photometry (Table 5). This may indicate that a substantial number of the objects in the original Dickinson et al. (2004) compilation were simply spurious sources. A cursory examination of these sources in the current ACS GOODS reduction bears out this supposition. From the HUDF, BSEM make the point that the cumulative surface density of i-dropouts is only 0.1±0.1 arcmin$^{-2}$ to $z_{850,AB}$ $\sim$ 26.5. While the present results roughly corroborate this claim, we find a slightly higher density (0.26 arcmin$^{-2}$) to the same bright limit in our corrected counts (Table 10). The current value is a bit lower than the completeness corrected 0.5 ± 0.2 i-dropouts arcmin$^{-2}$ cited in our earlier study on the RDCS1252-2927 + HDF-N fields (Bouwens et al. 2003b), but this appears to have been the result of large scale structure (B05b) and lensing by the prominent foreground cluster in that study. This surface density (0.26 arcmin$^{-2}$) also appears to be consistent with the 3-epoch estimate from the GOODS team, if we assume the 75% contamination fraction derived earlier (and apply a small completeness correction).

4.2. Is the Surface Density of i-dropouts in the UDF Typical?

The normalization of the i-dropout counts in a given field can show large variations (e.g., 35% RMS for a single ACS field) due to large scale structure (“cosmic variance”). In §3.6, we were able to estimate the normalizations for our fields relative to the large area GOODS fields. One field that was of particular concern in this analysis was the UDF because (1) it provides our best constraint on the number of faint i-dropouts and (2) it was selected to contain one particularly bright $z_{850,AB}$ = 25.0 i-dropout. Since rare objects are typically associated with overdensities, one might have expected the i-dropouts in the UDF to be overdense relative to the cosmic average, compromising any LF we might have determined using its data.

In §3.6, we showed that this is not likely an important concern, and that i-dropouts in the UDF have a surface density which is just 0.76 ± 0.20× that of the two GOODS fields (and thus the UDF may even be underdense relative to the cosmic average). Nevertheless, one might have expected this to be a concern given the recent findings by Malhotra et al. (2005) using the UDF GRISM data. Comparing the redshift distribution of i-dropouts they observed with that obtained from their modelling, Malhotra et al. (2005) argued that the UDF contained a $\sim 2 \times$ overdensity in the number of i-dropouts at $z = 5.9 \pm 0.2$ (15 of the total 23 i-dropouts). At first glance, these results may seem contradictory to our own, but one needs to remember that the Malhotra et al. (2005) measurement is really just a comparison between the volume density of i-dropouts inside the interval $z \sim 5.9 \pm 0.2$ and that outside it. Since the comparison was made entirely within the area of the UDF, it simply provides us with information on the large-scale structure at $z \sim 6$ along that line of sight.

5. LUMINOSITY FUNCTION

The combined data from the UDF, UDF-Ps, and GOODS fields provide a unique opportunity to derive...
the luminosity function at $z \sim 6$ to unprecedented depths and accuracy. Such detail is important for making accurate inferences about galaxy evolution and the reionization of the universe. It allows us to address questions about the subsequent evolution of $UV$-bright galaxies to $z \sim 3$, indicating whether there has been evolution in $L_\ast$, $\phi_\ast$, or $\alpha$. It also allows us to make reliable estimates of the UV background produced by $z \sim 6$ galaxies. The UV background density is crucial for assessing the impact of $z \sim 6$ galaxies on reionization.

Estimating the LF would be straightforward if there was a simple way of converting the observed fluxes $m$ to an absolute magnitude $M$ and moreover are not available for many of our fields. By contrast, our infrared fluxes are attenuated by the forest and thus conversions to absolute $UV$-binty:

$\phi(M) = \sum_k \phi_k W(M - M_k)$ (2)

where $\phi_k$ is the selection function, $M$ is the apparent $z_{580}$-band magnitude, $\phi_\ast$ is the selection function, and $M$ is the absolute magnitude at 1350 $\AA$. The absolute magnitude $M$ is a function of both the apparent magnitude $m$ and redshift $z$.

The selection function $P(m,z)$ can be estimated by projecting a complete $B$-dropout sample from the UDF (Bouwens et al. 2004b) to $z \sim 5-7$ and rescaling it using a similar procedure to that described in §3.1. For this projection, a $(1+z)^{-1.3}$ size scaling (for fixed luminosity: §3.7) is adopted. The projected $B$-dropout sample is assumed to have a $UV$-continuum slope $\beta$ with mean $-1.8$ and $1\sigma$ scatter of 0.5, similar to our fits in §3.4. Motivated by the findings of Stanway et al. (2004a) and Dow-Hygelund et al. (2005), we also assume that 25% of the projected $B$-dropouts have Ly$\alpha$ emission with an equivalent width of 30 $\AA$. This latter assumption provides a rough account for the bias introduced by the current $(t_{775} - z_{580})_{AB} > 1.3$ selection against galaxies with strong Ly$\alpha$ emission at $z \sim 5.5 - 5.9$ (Malhotra et al. 2005; Figure 6 of Dow-Hygelund et al. 2005). Since Ly$\alpha$ emission falls in the $t_{775}$-band for objects at these redshifts, such objects will not readily show up as dropouts. This reduces the selection volume for $z \sim 6$ galaxies by $\sim 3\%$. The selection function we derive is shown in Figure 9.

5.1. Direct Method

In this section, we present our primary determination of the rest-frame UV LF at $z \sim 6$. We express the LF in terms of a set of stepwise functions $\phi_k W(M - M_k)$ of half-magnitude width:

$\phi(M) = \sum_k \phi_k W(M - M_k)$ (2)

where

$W(x) = \begin{cases} 0, & x < -1/4 \\ 1 - 4/3x, & -1/4 < x < 1/4 \\ 1, & x > 1/4 \end{cases}$ (3)

We then derive the coefficients on the stepwise function through a maximum likelihood procedure, from a fit to the observed counts (Table 10). To simplify the computation, we derive kernels $V_{m,k}$:

$V_{m,k} = \int_z \int_{m-1/4}^{m+1/4} W(M'(z) - M_k) P(m', z) \frac{dV}{dz} dm'dz$ (4)

With this definition, Eq. (1) reduces to

$\sum_k \phi_k V_{m,k} = N_m$ (5)

where $N_m = \int_{m-1/4}^{m+1/4} N(m') dm'$. One example of the $V_{m,k}$ kernel is shown in Figure 8. Since our procedure here is essentially a deconvolution of $N_m$ (to obtain $\phi_k$), the LF we derive will contain more noise than the original counts. As a result (and because of the Poissonian noise in the observed counts at $z_{580,AB} > 27.5$), we have enlarged the size of our faintest two bins ($M_{1350,AB} > -19$) to be 1.0 mag in width.

The resultant LF is shown in Figure 11 (see also Table 11) and extends over 2 orders of magnitude: from $4 L_{z=3}^\ast$ down to 0.04 $L_{z=3}^\ast$. Remarkably, this is fainter than what Steidel et al. (1999) was able to obtain at $z \sim 3$ (where the limit was $\approx 0.1 L_{z=3}^\ast$). As a check on

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{figure8.png}
\caption{$z_{580}$-band magnitude vs. redshift (solid line) for objects of a fixed luminosity (here a $L_{z=3}^\ast$ galaxy). Consequently, objects at a particular $z_{580}$-band magnitude can correspond to a wide range of luminosities (e.g., a $z_{580,AB} \sim 27$ $i$-dropout would correspond to a 0.3$L_{z=3}^\ast$ object at $z \sim 5.5$ and a 2.5$L_{z=3}^\ast$ object at $z \sim 7$). To cope with this issue, we model the redshift distribution and convolve the LF ($\phi_\ast$) over the relevant selection volume when fitting the observed counts $N_m$ (Eq. 5). One example of the effective kernel $V_{m,k}$ (Eq. 4) used in these convolutions is shown here in the inset (for an object whose absolute magnitude corresponds to $L_{z=3}^\ast$). The effective kernels for other absolute magnitudes are similar. The vertical axis for the inset is in units of the selection volume per unit area per unit magnitude (Mpc$^3$/arcmin$^2$/mag).
\end{figure}
the current procedure, we repeated it on the surface density predictions made in Figure 6 (lowest panel) based on the $z \sim 3$ LF (Steidel et al. 1999) and were able to recover the input LF. For context, we present the predicted redshift distribution for this LF (and our UDF i-dropout selection) in Figure 10.

In addition to breaking up the LF in stepwise intervals, it has also become conventional to parametrize it in terms of a Schechter function (Schechter 1976). Because of the degeneracies amongst the parameters $\alpha$, $\phi^*$, and $M^*$, the results are expressed as likelihood contours (solid blue lines: Figure 12). In deriving these contours, it was reasonable to allow $\alpha$ to extend to values as steep as $-2$ since the LF must cut off at some physical scale (and therefore the luminosity density, although formally divergent for $\alpha < -2$, will converge). We evaluate the likelihood of different Schechter parametrizations by calculating the equivalent $\phi_k$’s for the parametrization and then comparing against the observed counts $N_m$ (Table 10) using Eq. (5). To illustrate the effect of fixing the different Schechter parameters from $z \sim 3$, solid green contours are overplotted in Figure 12. The present results can be put in context by comparing against the equivalent $z \sim 3$ determinations (Steidel et al. 1999). In the leftmost two plots, we see evidence for lower characteristic luminosities $M^*$ at $z \sim 6$, with little change in $\phi^*$ or $\alpha$. Fainter values of $M^*$ are favored at 82% confidence. In the rightmost plot, we can see that these results are also consistent with a steeper value for $\alpha$ at $z \sim 6$, even if no change in $M^*$ is allowed. However, scenarios, such as density evolution ($\phi^*$) which do not include these changes in $M^*$ (towards fainter values) or $\alpha$ (towards steeper values) are excluded at 99.9999% confidence. Our most likely values for $\phi^*$, $M_{1350,AB}$, and $\alpha$ are $1.76^{+1.66}_{-0.71} \times 10^{-3}$ Mpc$^{-3}$, $-20.20 \pm 0.35$, and $-1.74 \pm 0.24$, respectively. As illustrated in Figure 11 (solid line), this fit is in good agreement with the stepwise LF determined earlier. Because of the proximity of the present faint-end slope to $-2$, where the integral of the total light diverges, extrapolations to zero luminosity can be somewhat uncertain. A much more robust number is the total luminosity density integrated down to the approximate faint-end limit of the UDF $(0.04L^*_z=3)$: $1.46\pm0.14\times10^{26}$ ergs s$^{-1}$ Hz$^{-1}$ Mpc$^{-3}$. This is equal to 0.68×, 0.50×, and 0.24× the luminosity density integrated down to zero assuming faint-end slopes $\alpha$ of $-1.6$, $-1.7$, and $-1.9$, respectively.

### 5.2. STY Method

A more conventional way of deriving the LF (across multiple fields) is to use the Sandage, Tammann, & Yahil (1999: STY) fitting procedure. This procedure has the advantage of being relatively insensitive to large-scale structure. Only the shape of the luminosity function factors into the fits and not the normalization, allowing one to derive extremely robust measures on the overall shape. We do not use this procedure as our primary fitting procedure since our degradation procedure (§3.6) provides us with a slightly more direct measure of the field-to-field variance (the STY approach may be more sensitive to errors in our transfer functions). However, as the reader will see, the results are in very good agreement, suggesting that our basic results here are robust.

As with our primary approach, an important complication is the rather inexact relationship between apparent and absolute magnitudes (Figure 8). This makes it more convenient to work in terms of the apparent rather than absolute magnitudes; and our procedure becomes one where we are maximizing the likelihood of producing the observed counts (here distributed over three different fields) given a LF. In detail, this approach really is not that different from what we performed in §3.6 to match up the counts from our three different data sets, and so it should not be surprising that the best-fit parameters we obtained from this procedure, i.e., $M_{1350,AB} = -20.19$

<table>
<thead>
<tr>
<th>$M_{1350,AB}$</th>
<th>$\phi_k$ (Mpc$^{-3}$ mag$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$-21.943$</td>
<td>$0.00001 \pm 0.00001$</td>
</tr>
<tr>
<td>$-21.443$</td>
<td>$0.00005 \pm 0.00003$</td>
</tr>
<tr>
<td>$-20.943$</td>
<td>$0.00008 \pm 0.00006$</td>
</tr>
<tr>
<td>$-20.443$</td>
<td>$0.00026 \pm 0.00010$</td>
</tr>
<tr>
<td>$-19.943$</td>
<td>$0.00105 \pm 0.00025$</td>
</tr>
<tr>
<td>$-19.443$</td>
<td>$0.00250 \pm 0.00094$</td>
</tr>
<tr>
<td>$-18.693$</td>
<td>$0.00289 \pm 0.00095$</td>
</tr>
<tr>
<td>$-17.693$</td>
<td>$0.00677 \pm 0.00182$</td>
</tr>
</tbody>
</table>

### Fig. 9—The probability $P(m, z)$ that some object of apparent $z_{580,AB}$-band magnitude and redshift $z$ is included in our UDF i-dropout sample. This function was computed by projecting a UDF $B_{435}$-dropout sample (Bouwens et al. 2004b) to $z \sim 6$ assuming a $(1 + z)^{-1.1}$ size scaling (for fixed luminosity: §3.7). Other scalings (e.g., $(1 + z)^{-1}$ or $(1 + z)^{-1.5}$) yield only modest differences with respect to the adopted selection function $P(m, z)$ and therefore only have a minor effect on the shape of the LF (e.g., $\Delta \phi = \pm 0.1$). The rest-frame UV slopes $\beta$ of our input sample are assumed to have a mean of $-1.8$, with a 1σ scatter of $0.5$. 25% of the objects are assumed to have a Lyα equivalent width of $30$ Å (Dow-Hygelund et al. 2005). This function does not include the small incompleteness (≈11–17%) due to blending with foreground galaxies (Appendix D.1).
and $\alpha = -1.69$, and their likelihood contours (Figure 13) are in good agreement with those obtained with our primary methodology (Figure 12). The $\phi^*$ we derive fixing the shape of the LF and fitting to the number counts in the two GOODS fields (middle panel of Figure 6) is $1.86 \times 10^{-3}$ Mpc$^{-3}$ and also quite consistent.

5.3. Direct Method (without LSS correction)

Finally, it is interesting to compute the $z \sim 6$ LF but without any correction for large-scale structure (“cosmic variance”). Since field-to-field variations (i.e., 35% RMS for a single ACS: §3.6) are only slightly larger than our measurement errors on these variations (the uncertainties on the normalization factors for the UDF are 26% RMS; see Table 9), the LF we derive ignoring the normalization altogether (i.e., assuming each field is representative of the cosmic average) will be fairly competitive with our primary determination (§5.1). Meanwhile, differences we observe relative to this determination can provide us with a good sense for the representative errors. Reducing the LF with these assumptions, we obtained the following best-fit Schechter parameters: $\phi^* = 1.50 \times 10^{-3}$ Mpc$^{-3}$, $M_{1350,AB} = -20.23$, $\alpha = -1.62$. Encouragingly, these values are only slightly different from those obtained from the two previous methods (Figures 12 and 13), and in fact, we might have expected this level of agreement from some simulations we ran to assess the impact of cosmic variance on the derived LF (Appendix E).

5.4. Luminosity Densities

Having obtained a basic fit to the observed LF, we can move on to look at the UV continuum luminosity density and how it compares with previous determinations at higher and lower redshift. Because of the limited sensitivities of the highest redshift probes (e.g., the Bouwens et al. 2004c study at $z \approx 7$–8 and the Bouwens et al. 2005a study at $z \approx 10$), we make these comparisons down to two different luminosity limits: 0.3 $\times$ and 0.04 $\times$ the characteristic luminosity at $z = 3$ (Steidel et al. 1999). This is important to properly account for possible evolution in the characteristic luminosity $L^*$ or faint-end slope $\alpha$ with redshift. To a limiting magnitude of $0.3L^*_z$, the present LF integrates out to $4.7 \pm 0.6 \times 10^{25}$ ergs s$^{-1}$ Hz$^{-1}$ Mpc$^{-3}$.

To convert these UV luminosity densities into star formation rate densities (uncorrected for extinction), we assume a Salpeter IMF and use the now somewhat canonical conversion factors of Madau et al. (1998):

$$L_{UV} = \text{const} \times \frac{\text{SFR}}{M_\odot \text{yr}^{-1}} \text{ergs s}^{-1} \text{Hz}^{-1} \text{Mpc}^{-3}$$

(6)

where const $= 8.0 \times 10^{27}$ at 1500 Å. Both the present luminosity densities and star formation rates are shown in Figure 14 relative to many previous determinations (Steidel et al. 1999; Giavalisco et al. 2004b; Bouwens et al. 2004a; BSEM; Bouwens et al. 2004c; Schiminovich et al. 2005; Bouwens et al. 2005a). The fall off in the luminosity density towards high redshift is much sharper at brighter luminosities ($> 0.3L^*_z$) ($\rho(z = 6)/\rho(z =$
3) = 0.43 ± 0.05) than it is when integrated down to
0.04L_{z=3}^∗(\rho(z=6))/\rho(z=3) = 0.68 ± 0.08).

6. DISCUSSION

The combination of the UDF, UDF-Ps, and GOODS datasets, especially the very deep UDF data, provides a unique opportunity to explore a number of issues for \( z \sim 6 \) galaxies. These include refining our knowledge of the rest-frame UV-continuum luminosity function, assessing the impact of \( z \sim 6 \) galaxies on the reionization of the universe, and using \( z \sim 6 \) as a baseline for assessing evolution to even higher redshift. Our analysis of the rest-frame UV-colors also permits us to revisit the issue of a possible evolution in the UV-continuum slope \( \beta \).

6.1. UV Continuum Luminosity Function

One of the principal goals of this paper was to obtain a fairly optimal determination of the luminosity function in the rest-frame continuum UV (\( \sim 1350\AA \)). The present approach has several important advantages over several previous derivations (Dickinson et al. 2004; Yan & Windhorst 2004a; Bouwens et al. 2004a; BSEM; Yan & Windhorst 2004b; Malhotra et al. 2005). These include obtaining a self-consistent selection of \( i \)-dropouts from three of the deepest, widest area data sets (GOODS, UDF-Ps, UDF); systematic use of the deeper ACS and infrared data to derive completeness, flux, and contamination corrections; use of the average UV continuum colors in our selection volume estimates; an inclusion of the selection biases against strong Ly\( \alpha \) emitters in these same selection volume estimates; and a detailed matching-up of the surface density of \( i \)-dropouts in our deeper fields with that obtained in shallower, wider area fields to ensure a proper normalization of the overall LF.

The current refinement to the \( z \sim 6 \) LF puts us in a good position to examine several previous determinations of this LF and the associated claims for evolution from \( z \sim 3 \) (Dickinson et al. 2004; Yan & Windhorst 2004a; Bouwens et al. 2004a; BSEM; Yan & Windhorst 2004b; Malhotra et al. 2005). A summary of many previous Schechter parametrizations are given in Table 12 and plotted relative to the current determination in Figure 15. We break this discussion between the bright and faint ends of the LF. At the bright end (\( M_{1350,AB}^* \approx -21 \)), we find a substantial \( \sim 6 \times \) deficit relative to the \( z \sim 3 \) LF. This supports the initial findings of Stanway et al. (2003), Dickinson et al. (2004), and Stanway et al. (2004b). Our current estimate for the number density of \( i \)-dropouts at the bright end is slightly smaller than what we reported in two previous studies (Bouwens et al. 2003b; Bouwens et al. 2004a). In the first case this was because of a substantial (\( \approx 2 \times \)) overdensity in the RDCS1252-2927 field relative to the cosmic average (§4.1; B05b) and in the second case it was because of slight (\( \sim 20\% \)) overestimates of the surface densities and completeness present in the GOODS fields (Bouwens et al. 2004a). The number density is also less than re-
reported by Yan & Windhorst (2004). This appears to have been due to their reliance on the 3-epoch GOODS i-dropout catalog (Dickinson et al. 2004) which, as we discussed earlier (§4.1), overestimated the surface density of i-dropouts. Recent searches at bright magnitudes ($z_{AB, r} < 25.6$) with Subaru also find strong ($\approx 11\times$) deficits (Shimasaku et al. 2005).

At fainter luminosities, the $z \sim 6$ LF shows much better agreement with $z \sim 3$ than at the bright end. This suggests evolution. As discussed earlier (§5), the simplest way to accommodate these changes is through an evolution of the characteristic luminosity ($82\%$ confidence). Our best-fit is $0.7 \pm 0.4$ mag brightening in $M^*$. An evolution of the faint-end slope $\alpha$ to $-1.9$ can also help (from $-1.6$ at $z \sim 3$: Steidel et al. 1999). The latter option echoes earlier claims made by Yan & Windhorst (2004b) for a steep faint-end slope ($\alpha = -1.8$) using data from the UDF. However such faint-end slopes do not appear to be required (Figure 12). The faint-end slope is nevertheless steeper than the $\alpha = -1.15$ determined in our earlier work using the UDF-Ps (Bouwens et al. 2004a). The shallower slope from that study appears to have derived from the significantly lower surface density of i-dropouts present in the UDF-Ps ($\sim 0.6\times$ the cosmic average; see Table 9). Contrary to the present work (§3.6), no attempt was made there to treat possible field-to-field variations and therefore the shape of the LF was affected. The Dickinson et al. (2004) determination, by contrast, was too high at lower luminosities. This appears to have been a consequence of their substantial underestimate of the contamination rate (§4.1).

The present determination also differs substantially from the best-fit LF of BSEM (Figure 15), particularly at the faint end where our LF is nearly a factor of $\sim 6$.
higher. Since the derived counts from BSEM agree fairly well with those of the present study (Figure 6), how can the differences in the LF be so large? The volume element does not appear to be the culprit since the BSEM no-evolution predictions from \( z \approx 3 \) (Steidel et al. 1999) closely match our own. The only possible explanation appears to be due to some peculiarity in the way that BSEM derived their best-fit parameters. From their figures 10 and 11, it would appear that BSEM conducted their fits (\( \chi^2 \)) on the cumulative counts, not the differential counts. If so, this would not be appropriate as the data points in the cumulative counts are not independent. Our own fits to their differential counts (blue circles in Figure 6) yield \( \phi_{150,AB}^* = -20.49 \) and \( \phi^* = 0.00097 \) assuming a fixed \( \alpha = -1.6 \). This fit gives a cumulative luminosity density to their faint-end limit \( (z_{150,AB} = 28.5) \) which is \( \approx 2.7 \times \) higher than their optimal fit (\( \alpha = 6 \times \) drop in \( \phi^* \) from \( z \approx 3 \)).

The BSEM work excepted, there has been a growing consensus among \( z \approx 6 \) studies that the evolution in the UV LF at high redshift occurs primarily at the bright end. Shimasaku et al. (2005) made a similar argument based upon a comparison of their bright i-dropout search with those obtained at fainter magnitudes (Bouwens et al. 2004a; BSEM; Yan & Windhorst 2004b). Such luminosity-dependent trends would also partially explain the supposed discrepancy (e.g., Trimble & Schwaned 2005; Stanway et al. 2004b) between several early \( z \approx 6 \) results, where different evolutionary factors were quoted relative to no-evolution \( z \approx 3 \) expectations (e.g., \( \approx 6 \) by Stanway et al. 2003 vs. \( \approx 2 \) by Bouwens et al. 2003b). Though it was once thought these differences might be due to uncertainties in the completeness and contamination rates (Bouwens et al. 2003b; Stanway et al. 2004b), it now appears that differences in the flux limit may have played an equally important role.\(^5\)

It seems relevant to step back and look at the observed evolution in the larger context of galaxy evolution. What is remarkable about the evolution we observe is that the characteristic luminosity of galaxies in the UV shows a significant increase over the range \( z \approx 6 \) to \( z \approx 3 \). This is in contrast to the strong decrease observed from \( z \approx 2 \) to \( z \approx 0 \) (Arnouts et al. 2005; Gabasch et al. 2004) and suggests that galaxy formation is a very different process early on than it is at much later times. At early times, it seems reasonable to imagine that this increase in luminosity we observe is just a simple consequence of the merging and coalescence of galaxies expected in hierarchical scenarios. The fact that this does not occur at later times suggests that something must halt this growth and even turn it around. Though we discuss it no further, two promising explanations for this turn-around include AGN feedback (e.g., Scannapieco et al. 2005; Scannapieco & Oh 2004; Binney 2004; Di Matteo et al. 2005) and the transition from cold to hot flows (e.g., Birnboim & Dekel 2003).

\(^5\) In principle, comparisons between the UV LF at \( z \approx 4 \) and \( z \approx 3 \) also inform our understanding of high-redshift galaxy evolution. Unfortunately, studies have come to different conclusions. Iwata et al. (2003) at \( z \approx 5 \) and Sawicki & Thompson (2005) at \( z \approx 4 \) report more evolution at the faint end of the LF relative to \( z \approx 3 \), while Ouchi et al. (2004) find more evolution at the bright end.

In light of the likely relationship between the luminosity evolution observed and the evolution of the mass function, it makes sense to examine this connection briefly. Figure 16 presents the mass function at \( z \approx 3 \) and \( z \approx 6 \) calculated from the Sheth & Tormen (1999) formalism and a \( \Lambda \)CDM power spectrum (Bardeen et al. 1986) with \( \sigma_8 = 0.9 \), \( \Omega_0 = 0.048 \), \( \Omega_M = 0.3 \), \( \Omega_\Lambda = 0.7 \), and \( H_0 = 70 \) km/s/Mpc. The horizontal blue arrow provides the likely mass range for i-dropouts in our sample (e.g., Cooray 2005). Besides an obvious evolution towards higher masses at later times (factor of \( \approx 3 \) change), the mass function is also expected to flatten (\( \Delta \alpha = +0.27 \) from \( z \approx 6 \) to \( z \approx 3 \)).

![Figure 16](https://example.com/fig16.png)

**Figure 16.** The mass function (comoving volume density) at \( z \approx 3 \) (dotted line) and \( z \approx 6 \) (solid line) calculated using the Sheth & Tormen (1999) formalism and a \( \Lambda \)CDM power spectrum (Bardeen et al. 1986) with \( \sigma_8 = 0.9 \), \( \Omega_0 = 0.048 \), \( \Omega_M = 0.3 \), \( \Omega_\Lambda = 0.7 \), and \( H_0 = 70 \) km/s/Mpc. The horizontal blue arrow provides the likely mass range for i-dropouts in our sample (e.g., Cooray 2005).

6.2. Rest-frame UV colors

The present sample also allowed us to place constraints on the mean redshift and rest-frame UV slope \( \beta \). We obtained these constraints using the measured optical-infrared colors for specific 1775-dropouts from the HUDF (Table 3). A comparison of our measured colors with
those obtained in two previous studies (Stanway et al. 2005; Yan & Windhorst 2004b) shows no large systematic differences, but considerable scatter ($\pm 0.15$ mag) for individual objects. The scatter becomes even larger ($>0.4$ mag) in cases of possible blending with foreground objects. Relative to previous measurements, we would expect our measurements to represent a modest improvement given our use of more optimized scalable apertures (thus avoiding most blending problems) and careful aperture corrections.

Despite no large systematics relative to previous measurements of the colors, the mean $\beta$ inferred in this study is $-1.8$, which is redder than the $\beta = -2.2$ inferred in the Stanway et al. (2005) study based on the same data. The principal reason for the difference here is that current inferences are based upon the $J - H$ colors while previous inferences were based upon the $z - J$ colors. Since the $z - J$ colors are highly influenced by the redshift of a source and moreover can be quite insensitive to rest-frame UV color (see B05b), the $J - H$ colors are a much more robust indicator of the rest-frame UV slope. $z - J$ colors are also more sensitive to errors in alignment, errors in the aperture corrections, and uncertainties in the optical to infrared zeropoints. Therefore, we consider the present determination to be an improvement on the Stanway et al. (2005) estimate (though current uncertainties in the zeropoints may make all present measures somewhat uncertain, i.e., $\Delta \beta = \pm 0.2$: §2.1).

Irrespective of the exact $\beta$, the mean rest-frame UV slope observed at $z \sim 6$ is bluer than that observed at $z \sim 3$. This evolution is consistent with a number of studies which have come out recently (Lehnert & Bremer al. 2003; Kneib et al. 2004; B05b; Bouwens et al. 2004c; Schaerer & Pelló 2005; Yan et al. 2005; cf. Ouchi et al. 2004) and point towards a lower mean dust extinction at higher redshift. Changes in age, metallicity, and the IMF have a much smaller effect on the rest-frame UV slope (e.g., Schaerer, 2003; Leitherer al. 1999). Moreover, a significant contribution from $\text{Ly} \alpha$ to the $J_{110}$ flux seems unlikely given constraints from emission line searches at $z \sim 6$ (Ajiki et al. 2003; Kodaira et al. 2003; Hu et al. 2004; BSEM; Stanway et al. 2004a; Nagao et al. 2004; Dow-Hygelund et al. 2005). This leaves an evolution in the dust content as the most natural way of accommodating this change (see also discussion in B05b).

One obvious consequence of this lower dust extinction is an evolution in the correction factor applied to the star formation rate densities inferred directly from the UV luminosity function (see also B05b; B05c; Stanway et al. 2004b). A convenient way of estimating the effect of this change is through the Meurer et al. (1999) fit relating the extinction $A_{1600}$ to the UV slope $\beta$: $A_{1600} = 4.43 + 1.99 \beta$. Though there is some uncertainty in the exact value of $\beta$ at $z \sim 6$ (and $z \sim 3$), it is useful to adopt some fiducial value of $\beta$ to estimate the size of the effect. Taking $\beta$ to equal $-1.8$ at $z \sim 6$ and $-1.5$ at $z \sim 3$ (Adelberger & Steidel 2000) yields an evolution of $\sim 2 \times$ in the attenuation factor at 1600Å, from $\sim 2$ at $z \sim 6$ to $\sim 4$ at $z \sim 3$. Since this is the same direction as the evolution of the UV LF, it appears that the real evolution (after correction for extinction) may be large indeed. So, instead of the $\sim 2 \times$ evolution in characteristic luminosities from $z \sim 6$ to $z \sim 3$, the real evolution in this quantity may be as large as $\sim 4 \times$ (after correction for extinction). It may also suggest that the total star formation rate (and UV luminosity) density at $z \sim 6$ (after correction for extinction and integrated to $0.04L_{\ast, z=3}$) is just $\sim 0.3 \times$ the value at $z \sim 3$ (instead of the $0.68 \times$ factor given in §5.4). We have included a simple illustration of this effect in Figure 17 using several of the more representative determinations of the UV luminosity density (Steidel et al. 1999; Schiminovich et al. 2005). The dust corrections we have applied here are $\sim 2 \times (0.8 \text{ mag})$ at $z \sim 6$ and from Schiminovich et al. (2005) otherwise. As is apparent from the figure, such changes have wide-range implications and indicate a much more rapid rise in the star formation history from $z \sim 6$ to $z \sim 3$. Clearly, it will be important to confirm this change with other methods (e.g., by using stacked X-ray fluxes: Reddy & Steidel 2004).

Perhaps we should not be surprised by this evolution in the rest-frame UV slope or the dust extinction. Given the strong correlation between the total star formation rates and the dust extinction (Wang & Heckman 1996; Martin et al. 2005; Adelberger & Steidel 2000), we might have expected the extinction to be lower at the highest redshifts. The mass scales are expected to be lower there, and as we have observed, so are the typical UV luminosities and apparent star formation rates.

### 6.3. Reionization of the Universe

In light of the observational evidence that $z \sim 6$ marks the end of the reionization epoch (Becker et al. 2001; Fan
et al. 2002; White et al. 2003), it has become common to use the observed i-dropouts to comment on the possible reionization of the universe by photons arising from galaxies (e.g., Stanway et al. 2003; Lehner & Bremer 2003; Bouwens et al. 2003b; Giavalisco et al. 2004b; Dickinson et al. 2004; Stanway et al. 2004b; BSEM; Stiavelli et al. 2004b). An estimate of the star formation rate necessary to produce this reionizing flux can be made using the convenient formulation of Madau et al. (1999) modified to match the baryon density derived from the recent WMAP results (Bennett et al. 2003) and shifted to $z \sim 6$ (Bouwens et al. 2003; BSEM):

$$\rho_\star \approx (0.052 \, M_\odot \, yr^{-1} \, Mpc^{-3}) \left( \frac{0.5}{f_{\text{esc,rel}}} \right) C_{30} \left( \frac{1 + z}{7} \right)^3.$$  

(7)

where $\rho_\star$ is the star formation rate density, $C_{30}$ is the H i concentration factor $\langle \rho_{H\,i}^2 \rangle / \rho_{H\,i}^2 / 30$, and $f_{\text{esc,rel}}$ is the relative fraction of ionizing radiation escaping into the intergalactic medium to that escaping in the UV-continuum ($\sim 1500 \, \AA$). Unfortunately, current constraints on the total star formation rate $\rho_\star$ still remain poor. Though an integration of our best-fit LF to our faint-end limit and zero luminosity yields 0.018 $M_\odot \, yr^{-1} \, Mpc^{-3}$ and 0.037 $M_\odot \, yr^{-1} \, Mpc^{-3}$, respectively (somewhat smaller than the fiducial star formation rate needed), current constraints also allow for substantially steeper values of the faint-end slope (e.g., $\alpha \sim -1.9$: Figure 12), which would nearly double this value of $\rho_\star$ and hence be sufficient to reionize the universe in this formulation. Of course, it is definitely true that physical constraints become important at some faint-end slope (given limits on the total stellar mass or metals produced, e.g., Madau et al. 1998; Stiavelli et al. 2004a).

Despite current refinements to the $z \sim 6$ UV continuum LF, there continue to be substantial uncertainties in the role that $z \sim 6$ galaxies play in reionizing the universe. Indeed, we should not forget that we still do not have a direct measure of the ionizing radiation escaping into the IGM and are forced to rely on a proportionality factor, called the relative escape fraction, to convert the observed rest-frame continuum-UV flux into an ionizing flux. While most attempts to measure this escape fraction at $z \lesssim 3$ have thusfar only obtained upper limits (i.e., $< 0.1$ to $< 0.4$) (Leitherer et al. 1995; Hurwitz et al. 1997; Deharveng et al. 2001; Gilliengo et al. 2002; Fernández-Soto et al. 2003; Malkan et al. 2003; Inoue et al. 2005), there have been other notable efforts (e.g., Steidel et al. 2001) that have obtained much larger values ($\gtrsim 0.5$). The situation remains somewhat controversial. As a result of these and other uncertainties (e.g., Stiavelli et al. 2004b), there has been a wide range of different claims regarding the capacity of galaxies to reionize the universe. Some authors have claimed the observed galaxies aren’t sufficient to reionize the universe (BSEM) while others have claimed that they are, either because of a higher ionizing efficiency (Stiavelli et al. 2004b) or because of a large contribution from lower luminosity galaxies at the faint end of the luminosity function (Yan & Windhorst 2004a,b). This study (with its more detailed matching up of the different surveys) provides an important confirmation of this latter result – though it is not yet clear that the faint-end slope is especially steep (i.e., $\alpha \lesssim -1.8$: as argued by Yan & Windhorst 2004a,b). This being said, we would like to reemphasize the considerable uncertainties present at this stage and how little knowledge we have about how the escape fraction might behave, both in its redshift and luminosity dependence. Better constraints will be available when we are able (1) to better characterize the escape fraction and (2) to look at the ionizing flux of $z \gtrsim 3$ objects more directly (as one might obtain through proximity studies).

6.4. Implications for $z_{850}$-dropout Samples

Our redetermination of the $z \sim 6$ LF allows us to remark on recent $z \sim 7.5$ $z_{850}$-dropout samples selected from the HUDF. There are two recent samples worthwhile to comment on: the Yan & Windhorst (2004b) sample and the Bouwens et al. 2004c sample. Yan & Windhorst (2004b) performed a shallow search to $J_{110,AB} \sim 26.6$ and found only 1 candidate, which was just on the edge of their selection window. Since they predicted 2.9 candidates to a similar limit from their i-dropout LF assuming no-evolution, they interpreted this as a tentative indication for the onset of galaxy formation at $z \sim 6$. Performing a much deeper search ($H_{160,AB} < 27.5$) on the same field, Bouwens et al. (2004) found 5 such $z_{850}$-dropout candidates, 4 of which they assume to be real in their fiducial estimates (at least one was considered to be spurious due to the aggressive nature of the search). Comparing this with the 14 $z_{850}$-dropouts predicted assuming no-evolution from $z \sim 4$, this appeared consistent with a $\sim 3 - 5 \times$ drop in the number (luminosity density) of UV-bright objects from $z \sim 4$ to $z \sim 7.5$.

It is relevant to revisit these issues using the current $z \sim 6$ LF. Comparisons can be made by projecting the present i-dropout sample to $z \sim 6-9$ using our cloning machinery, adding it to the data and then reselecting it. In simulating the profiles of specific i-dropouts in our LF, we use scaled versions of specific i-dropouts from the UDF matched to the current i-dropout LF (Appendix F of B05b details an analogous modelling of U-dropouts using the HDF profiles). Running through this procedure, 0.7 $z_{850}$-dropouts are expected to $J_{110,AB} \sim 26.6$ vs. the 1 found, and 5.4 $z_{850}$-dropouts are expected to $H_{160,AB} \sim 27.5$ vs. the 4 fiducial candidates. This suggests that there has only been a modest increase in the number of bright objects from $z \sim 7-8$ to $z \sim 6$ though the uncertainties are large due to small number statistics, cosmic variance, and some questions about the $z_{850}$-dropout candidates themselves (items 2b, 2c, and 2f from Bouwens et al. 2004c).

7. CONCLUSIONS

We have compiled a sample of 506 i-dropouts ($z \sim 6$ galaxies) from the HUDF, the UDF–Parallel ACS fields (UDF–Ps), and the GOODS fields (368 arcmin$^2$), the latter enhanced by the ACS supernova search data (extending the depth of the ACS i and z-data by 0.2 and 0.4 mags, respectively – to be the v2.0 GOODS release). This statistically–robust sample consists of 122, 68, and 332 galaxies, respectively, from the three aforementioned fields and includes objects in the HUDF as faint as $z_{850,AB} \sim 29.5$ (16 of these i-dropouts are common between two of the above samples). The current sample of 506 galaxies represents the most comprehen-
Table 13: Properties of $z \sim 6$ Galaxies.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$r_{HI}$ (at $z_{5800,AB} \sim 27$)</td>
<td>$-0.8$ kpc ($\sim 0.14''$)</td>
</tr>
<tr>
<td>Size-Redshift Scaling</td>
<td>$\beta \sim 1.8 \pm 0.3$</td>
</tr>
<tr>
<td>UV slope $\beta$</td>
<td>$(1 + z) \sim 1.8 \pm 0.3$</td>
</tr>
<tr>
<td>M$^*_{1350,AB}$</td>
<td>$0.00176 \pm 0.00071$ Mpc$^{-3}$</td>
</tr>
<tr>
<td>$M^*_{1350,AB}$ $\alpha$</td>
<td>$-20.0 \pm 0.35$</td>
</tr>
<tr>
<td>$\mathcal{L}<em>{1350}(&gt;0.3L^*</em>{MIR})$</td>
<td>$4.7 \pm 0.6 \times 10^{25}$ ergs/s/Mpc$^3$</td>
</tr>
<tr>
<td>$\mathcal{L}<em>{1350}(&gt;0.04L^*</em>{MIR})$</td>
<td>$1.46 \pm 0.14 \times 10^{26}$ ergs/s/Mpc$^3$</td>
</tr>
</tbody>
</table>

We select these galaxies using the well-established dropout technique, with an i-dropout criterion ($[i'775 - z_{5800}]_{AB} > 1.3$, $(V_{606} - z_{5800})_{AB} > 2.8$) and demonstrate that the contamination levels on our selection are $\lesssim 8\%$ (i.e., $\leq 92\%$ are real: Appendix D.4).

Contamination is a potentially serious concern for dropout samples. We gave particular attention to four sources of contamination: intrinsically red–low-redshift galaxies, stars, spurious sources, and low-redshift galaxies scattering into the selection region (photometric scatter). We established the contamination levels by performing object selection in degraded versions of the deepest fields, and utilized the deep (NICMOS) and wide-area (ISAC) infrared images of these fields. As we discuss in Appendix D.4, red galaxies only appear to be a significant source of contamination ($18.5\%$ of our i-dropout candidates) at bright magnitudes ($25 < z_{5800,AB} < 26$) and again possibly at the faintest magnitudes ($10^{1-5}\%$: $z_{5800,AB} > 25$). Contamination from photometric scatter is also small ($<10\%$) and only important near the faint-end limit. Contamination from stars is uniformly low at all magnitudes ($<3\%$: after filtering out the few obvious bright cases), while that from spurious sources is insignificant. The present i-dropout catalogs are extremely clean (overall $\lesssim 8\%$ contamination).

An optimal measure of the i-dropout surface densities over a 5 magnitude range ($24.5 < z_{5800,AB} < 29.5$) is determined from our three samples (HUDF, UDF-Ps, enhanced GOODS). Detailed degradation experiments are made on our deeper data sets to understand object selection and photometry in our shallower fields and to derive completeness, flux and contamination corrections. These corrections are applied to establish a common baseline across all data sets. To remove the effects of large-scale structure (we expect $\sim 35\%$ RMS variations in the surface density of i-dropouts over 12 arcmin$^2$ ACS fields) when combining our three i-dropout samples, we carefully match up the surface density of i-dropouts in the deeper UDF and UDF-Ps probes to that found in the two enhanced wide-area GOODS fields ($\S 3.6$). The UDF and UDF-Ps fields are underdense ($0.76 \pm 0.20 \times$ and $0.66 \pm 0.16 \times$) relative to the cosmic average defined by the two GOODS fields (Table 9).

Finally, we use our derived surface densities to calculate a rest-frame UV LF at $z \sim 6$, and compare this LF with lower redshift ($z \sim 3$) LFs. Quantitative estimates of both the sizes and UV colors of objects in our samples are used to ensure rigor in the derived LF. Our principal findings are summarized in Table 13 and are as follows:

**Galaxy Sizes**: Typical i-dropouts at $z_{5800,AB} \sim 27$ (from the UDF-Ps and UDF) have PSF-corrected half-light radii of $\sim 0.8$ kpc or $\sim 0.14''$ (Figure 7: $\S 3.7$). Comparing the observed sizes of i-dropouts from our three different data sets with that predicted using different scalings of a $z \sim 2.5$ U-dropout sample (B05b), we make inferences about how the physical sizes of galaxies depend on redshift (for fixed luminosity). Our best-fit is a $(1 + z)^{-1.1 \pm 0.3}$ scaling, which is in good agreement with several previous determinations (Ferguson et al. 2004; Bouwens et al. 2004a,b; B05c).

**Rest-frame UV Colors/UV-to-total SFR Corrections**: Modelling the $z - J_{110}$ and $J_{110} - H_{160}$ colors of i-dropouts from the NICMOS HUDF (see $\S 3.3$), we construct models for the rest-frame UV colors of the i-dropout population and the redshift distribution (which we find peaks at $z \lesssim 6$ – see also Malhotra et al. 2005). The mean rest-frame UV spectral slope $\beta$ we infer is $-1.8$. This is bluer than the $-1.5$ observed at $z \sim 3$ (Adelberger & Steidel 2000), but redder than the Stanway et al. (2005) estimate of $\beta = -2.2$ at $z \sim 6$ using the same data. A similar evolution from bluer spectral slopes has already been noted in a number of other high-redshift ($z > 3$) studies (Lehnert & Bremer 2003; Kneib et al. 2004; B05b; Bouwens et al. 2004c; Schaerer & Pelló 2005; Yan et al. 2005; cf. Ouchi et al. 2004). The most natural explanation for this evolution is an increase in the dust content at later cosmic times. The most salient implication of such an evolution is the effect it would have on the inferred star formation rate densities. Using the Meurer et al. (1999) relation between the UV slope $\beta$ and the extinction $A_{1600}$, we estimate this factor (at 1600Å) to change from $\sim 2 \times$ at $z \sim 6$ to $\sim 4 \times$ at $z \sim 3$.

**Luminosity Function**: Using the surface densities of i-dropouts from our data sets and a computed selection function $P(m,z)$, we derive the rest-frame continuum UV ($z_{1350} \sim 6$) LF at $z \sim 6$ from $4L^*_{1350} \sim 3$ to $0.04L^*_{1350} = 17.5$ (M$_{1350,AB} \sim -17.5$: see $\S 5$). This is fainter than Steidel et al. (1999) was able to obtain at $z \sim 3$ ($0.1L^*_{1350}$). The likelihood parameters we derive for different Schechter parameterizations suggest that there has been an increase in the characteristic luminosity $M^*_{1350,AB}$ from $z \sim 6$ to $z \sim 3$ (82% confidence). The best-fit is a $0.7 \pm 0.4$ mag brightening. This evolution in $M^*$ can be partially offset by changes in the faint-end slope $\alpha$ (from $-1.9$ at $z \sim 6$ to $-1.6$ at $z \sim 3$; Steidel et al. 1999). Scenarios, such as density evolution ($\phi^*$), which do not include this evolution in $M^*$ or $\alpha$ are excluded at 99.9999% confidence. The best-fit Schechter parameters are $M^*_{1350,AB} = -20.20 \pm 0.35$, $\alpha = -1.74 \pm 0.24$, and $\phi^* = 1.76_{-0.71}^{+1.16} \times 10^{-3}$ Mpc$^{-3}$.

**Luminosity/Star Formation Rate Density**: The rest-frame continuum UV ($z_{1350} \sim 1350$) luminosity density at
$z \sim 6$ is $4.7 \pm 0.6 \times 10^{25} \text{ergs/s/Hz/Mpc}^3$ integrated down to $0.3L_\ast$ and $1.46 \pm 0.14 \times 10^{26} \text{ergs/s/Hz/Mpc}^3$ integrated down to $0.04L_\ast$ ($\S 5$). This is $0.43 \pm 0.05 \times$ and $0.68 \pm 0.08 \times$, respectively, the luminosity density at $z \sim 3$ (Steidel et al. 1999) to comparable faint-end limits. The large dispersion in previous results at $z \sim 6$ seems at least in part to have been due to a dependence on the adopted faint-end limit (e.g., compare the panels in Figure 14).

Adopting the evolution in the UV-to-total correction factors quoted earlier (and thus dust content as the reason for the change in the rest-frame UV colors), we infer a much stronger evolution in the star formation rate density over the range $z \sim 6$ to $z \sim 3$ than is found in the luminosity density. Using the Meurer et al. (1999) prescription, we estimate that the star formation rate density at $z \sim 6$ is only $\sim 0.3 \times$ that at $z \sim 3$ (to $0.04L_\ast$), quite different from the change in the luminosity density (where it is $0.68 \times$).

Reionization of the Universe: Assuming an escape fraction of 0.5 and H I concentration factor of 30, we estimate that a star formation rate density of $0.052 M_\odot \text{yr}^{-1} \text{Mpc}^{-3}$ is needed to reionize the universe using the Madau et al. (1999) formulation ($\S 6.3$). This is to be compared to the 0.018 $M_\odot \text{yr}^{-1} \text{Mpc}^{-3}$ observed to the limit of our probe and the 0.037 $M_\odot \text{yr}^{-1} \text{Mpc}^{-3}$ obtained by extrapolating our best-fit LF to zero luminosity. Despite being lower than the fiducial star formation rate densities required, there are sufficient uncertainties at present (particularly in the escape fraction, ionizing efficiency, and faint-end slope) that this factor of 2 difference is not significant. $z \sim 6$ galaxies seem capable of reionizing the universe (see also Stiavelli et al. 2004b).

$z \sim 7-8$ Galaxies: Projecting the present i-dropout LF to $z \sim 6-9$, we make an estimate of the number of $z_{850}$-dropouts that would have been found in a number of recent works (Yan & Windhorst 2004b; Bouwens et al. 2004c). We estimate $0.7 z_{850}$-dropouts to $J_{110, AB} \sim 26.6$ vs. the 1 found and $5.4 z_{850}$-dropouts to $H_{160, AB} \sim 27.5$ vs. the 4 fiducial candidates found. Despite substantial uncertainties, this suggests that the rest-frame UV LF only shows a slight decline from $z \sim 6$ to $z \sim 7.5$ ($\S 6.4$).

The HST ACS data from the UDF, the UDF Parallel fields and GOODS fields (enhanced by the extensive supernova search data) enabled us to detect 506 $z \sim 6$ galaxies. This remarkable sample has been used to derive a rest-frame UV luminosity function at $z \sim 6$ that extends three magnitudes below $L^*$ (to 0.04$L^*$), as well as providing improved constraints on size and color evolution – clearly establishing that galaxies are smaller and bluer. The $z \sim 6$ LF demonstrates that the brightest galaxies are less luminous at $z \sim 6$, i.e., that luminosity evolution is the dominant characteristic of the evolving galaxy population between from $z \sim 6$ (0.9 Gyr) to $z \sim 3$ (2 Gyr). The broad consistency of these results with the expectations of hierarchical models is encouraging. However, it is the quantitative constraints made possible with current data sets that are really important. Like $z \sim 3$, $z \sim 6$ seems destined to become an important reference point in our studies of galaxy evolution, marking as it does the end of the reionization epoch and providing a useful baseline for theoretical exploration to even earlier times.

We are appreciative to the many individuals who had contributed to our cloning software with their thoughts, ideas, or other suggestions. We acknowledge useful discussions with Brandon Allgood, Tom Broadhurst, Andy Bunker, Daniel Eisenstein, Akio Inoue, Sangeeta Malhotra, James Rhoads, Evan Scannapieco, Daniel Schaerer, and Jason Tumlinson. We are indebted to Adam Riess for use of the ACS images from his SNe search program over the two GOODS fields, allowing us to make a much deeper reduction. We thank Dan Magee for helping to install “apsis” on our computer systems, Wei Zheng for checking the infrared fluxes on our i-dropout candidates over the ACS parallel fields (UDF-Ps), Ruben Salvaterra for a careful reading of this manuscript, and the referee for many helpful comments. ACS was developed under NASA contract NAS5-32865, and this research was supported under NASA grant HST-GO09803.05-A and NAG5-7697.

REFERENCES

To determine the noise model for each of our images, we measure the RMS variance in apertures of different weight maps, which are expressed in units of the inverse variance (equal to what they would be without any correlation). Appendix C, and Appendix D, we used this procedure to quantify the effect of S/N on object selection and photometry.

APPENDIX

A. V − Z COLOR CUT

In Appendix C, we describe some degradation experiments we performed. The purpose of these experiments was to understand object selection and photometry in our lower S/N data. One key goal was to determine the extent to which lower redshift objects contaminated our lower S/N samples, but to do this, we needed to know which objects on our images were at low redshift and which objects were likely at high redshift. Fortunately, we were able to use the photometric information available from our deeper (undegraded) data to accomplish this. We settled upon a \((V - 5800)_{AB} = 2.8\) color cut (Figure A18). Objects whose colors were greater than 2.8 in the deeper data were classified as high-redshift objects and those with colors less than 2.8 were classified as lower redshift contaminants. Though ideally this criterion would have allowed us to eliminate all low-redshift interlopers while including objects near the lower end of the selection window in redshift (i.e., \(z \sim 5.5\)), a clean separation isn't entirely possible. Excluding all \(V - 5800\) colors greater than 2.8 (Figure A18). Therefore the \(z \sim 5.5\) would nominally require the cut to be 2.9 (though a consideration of the photometric scatter, typically \(\sim 0.4\) mag, suggests that 2.5 would be better). Therefore, it was necessary for us to settle on some compromise between 2.5 and 3.5. We chose 2.8.

Such a choice may not be effective in identifying all lower redshift interlopers, however, since some objects with \((S5800 - K_J)_{AB} > 1.6\) colors are likely to have \((V - 5800)_{AB}\) colors greater than 2.8 (Figure A18). Therefore the contamination rate we determine from our degradation experiments (Appendix D.4.2) may be an underestimate. Fortunately, this is not a concern, since the very same objects will be included in our estimate of the contamination rate from extremely red objects (Appendix D.4.1).

B. DEGRADATION PROCEDURE

At several points in our analysis, we found it convenient to degrade our deeper data down to some shallower S/N level. In §2.3, we used this procedure to obtain a uniform S/N level across the two GOODS fields, and in §3.4, §3.6, Appendix C, and Appendix D, we used this procedure to quantitatively the effect of S/N on object selection and photometry.

Before discussing our degradation procedure, it is helpful to provide some background on both our noise models and weight maps, which are expressed in units of the inverse variance (equal to what they would be without any correlation in noise). To determine the noise model for each of our images, we measure the RMS variance in apertures of different...
Fig. A18.— Motivation for our $(V_{606} - z_{850})_{AB} > 2.8$ cut used for selecting $i$-dropouts. The measured $(V_{606} - z_{850})_{AB} / (z_{850} - K_s)_{AB}$ colors for spectroscopically confirmed $i$-dropouts (filled squares: Dickinson et al. 2004 and Malhotra et al. 2005) are contrasted with those obtained from a set of lower redshift interlopers selected in the CDF-South GOODS ISAAC footprint (filled circles: Table C14). The latter objects were selected to have $(i_{775} - z_{850})_{AB}$ colors redder than 1.0 and $(z_{850} - J)_{AB}$ colors redder than 0.8. The model colors are shown for three different UV spectral slopes $\beta$ and three different low-z interlopers (Coleman et al. 1980). Redshifts are marked on the diagram alongside the tracks. The $(V_{606} - z_{850})_{AB} = 2.8$ color cut shown here (vertical line) is used to discriminate against low-z early types and all later spectral types which enter into our sample (see Appendix A). Early-types not caught by the $(V_{606} - z_{850})_{AB}$ cut will be included in our estimates of the contamination fraction using the $z_{850} - K_s$ colors of objects from the CDF South GOODS ISAAC and UDF NICMOS data (see Figure D21, Table D21, and Appendix D.4.1). Red $(i_{775} - z_{850})_{AB} > 1.3$ objects with $(z_{850} - K_s)_{AB} > 1.6$ colors (horizontal line) are included in this contamination fraction.

sizes, and then find an RMS noise level and noise kernel which reproduced the observed variation in RMS noise as a function of aperture size. The best fit noise levels were then used to scale the weight (inverse variance) maps provided with the UDF (Beckwith et al. 2005) or obtained from our “apsis” software (Blakeslee et al. 2003).

Our degradation procedure for individual images was as follows. Difference images were generated to reflect the difference between the weight (inverse variance) maps and the threshold (minimal) value for the weight (inverse variance), thus allowing us to account for pixel-by-pixel differences in the S/N. These difference images were used to generate noise images that were to be added to the real data (the noise was assumed to have a normal distribution). However, before adding these noise images to the data, they were first convolved by the noise kernel determined for the data. The weights on all degraded pixels were set to the value demanded by our S/N requirements.

C. DEGRADATION EXPERIMENTS

To assess the completeness, contamination rate, and flux measurements in our shallower fields relative to our deeper fields, we degraded our deeper fields (UDF and UDF-Ps) down to the depths of our shallower fields (UDF-Ps and GOODS) in a series of experiments. These degradations provide a very natural way of estimating the effect that photometric scatter has on both our selection and measurement process. Experiments included degrading the UDF down to the depth of the first UDF parallel (UDFP1), degrading the UDF down to the depth of the second UDF parallel (UDFP2), degrading the UDF down to the depth of the GOODS fields, degrading UDFP1 down to the depth of the GOODS fields, and degrading UDFP2 down to the depth of the GOODS fields. Each degradation was repeated 10 times to minimize the dependence upon any particular noise realization. To maximize realism, we ensured that the pixel-by-pixel weight maps of the degraded images were identical to that of the shallower fields. This was of particular interest for the UDF-Ps (§2.2) because the depth in these fields varies by $\sim 0.4$ mag across the field of view. $i$-dropouts were then selected using the selection criteria of our shallower fields. Our degradation procedure is detailed in Appendix B.

Objects obviously associated with the diffraction wings of stars or the wings of ellipticals were eliminated to mimic the selection procedure used for the main catalog (where similar spurious sources were eliminated). Objects selected by this procedure were divided into two categories: contaminants and real objects. Objects with $(V_{606} - z_{850})_{AB}$ colors less than 2.8 were classified as contaminants and objects with $(V_{606} - z_{850})_{AB}$ colors greater than 2.8 were classified as high-redshift objects (see Appendix A). In a few cases, where it was clear that the $V$-band photometry...
was contaminated by a nearby foreground object, we reclassified what would otherwise be labelled a contaminant as a $z \sim 6$ $i$-dropout. Despite some ambiguity regarding the exact split between the two categories, our results are not expected to depend on the exact split chosen. More stringent ($V_{606} - z_{850} \lesssim 26.8$) cuts will result in a higher contamination rate for the shallower field, but this will be offset by a lower selection volume for our sample as a whole.

D. CORRECTIONS APPLIED TO OUR DATA

This section describes the corrections we applied to the surface densities of $i$-dropouts derived from our shallower data to put them on a similar footing to our deeper UDF data. These corrections compensate for the greater incompleteness levels, flux biases, and contamination expected to be present in the shallower data.

We start off this section by looking at what can be said about the completeness levels and flux biases by degrading the available data. Though these issues will eventually be treated here with transfer functions (Appendix D.3), our initial analyses here will provide for some valuable benchmarks we can use to assess the validity of the transfer functions we determine.

D.1. Completeness Corrections

A generic consequence of S/N thresholds and standard detection algorithms is an overall incompleteness at faint magnitudes and large sizes. An illustration of this is provided in Figure D19 for the three fields under study, and it is immediately apparent that the distribution of $i$-dropouts in the UDF, UDF-Ps, and GOODS fields do not extend much to the upper-right of the three 50% completeness contours shown. As a result, significant incompleteness is not expected until at least $z_{850,AB} \sim 26.8$ in the GOODS fields, $z_{850,AB} \sim 27.5$ in the UDF-Ps fields, and $z_{850,AB} \sim 29$ in the HUDF.

Perhaps, the most model-independent way of estimating the incompleteness in our shallower fields relative to the UDF is to degrade the deeper data sets down to the same S/N as these shallower fields and then repeat the selection procedure. The very similar PSFs, pixel sizes, and passbands for all data sets considered here make this a very straightforward process. The deeper data also provide a fairly natural way of determining the fraction of objects on the degraded frames which are high redshift objects and the fraction which are likely contaminants or noise. These simulations are described in Appendix C.

Comparing the surface density of the sources recovered in the deeper images with that recovered at the shallower depths (while excluding those objects whose $V_{606}$-band fluxes indicate they might be contaminants, Appendix D.4.2), we are able to compute the completeness for the different fields under study. The results of the simulations are given in Tables D15-D16 and can put together to obtain an estimate of the completeness relative to the UDF.

An application of binomial statistics to the results of Table D15 enables a fairly straightforward determination of the magnitude-dependent completeness of the UDF-Ps relative to the UDF. While a similar procedure can be used to calculate the completeness of the GOODS probe relative to the UDF, tighter constraints can be obtained by using objects from both the UDF (Table D15) and UDF-Ps (Table D16). This takes advantage of the fact that the UDF-Ps are significantly more complete than the GOODS fields at all magnitudes. However, to use the results from the UDF-Ps, we need to make a small correction for the small differences in the completeness between the UDF and UDF-Ps selections (based upon the results in Table D15). The 1σ confidence intervals on the incompleteness of both fields are tabulated in Table D17.

Finally, we discuss issues of incompleteness due to blending with foreground galaxies (object overlap). Though not generally considered to be an important source of incompleteness ($\lesssim 10%$) for HST studies, here blending played a slightly larger role. This was due to our choice of blending parameters ($\text{DEBLEND\_MINCONT}=0.15$), which we adopted to ensure that SExtractor kept many of the more lumpy dropouts in our sample in a single piece (see §3).

To compute the incompleteness from blending, we included $i$-dropouts from our samples onto the image frames, and then attempted to recover them with our selection procedure. We used analytic versions (i.e., best-fit exponential
Fig. D19.— Size-magnitude diagram for $i$-dropout objects from the UDF (red dots), UDF-Ps (blue dots), and GOODS fields (black dots). The 50% completeness limits are overplotted for the three fields assuming an $r^{1/4}$ surface brightness profile. These limits were determined by laying down galaxies of different sizes and total magnitudes on a noise frame and then attempting to reselect it using our selection procedure (§3.1). Comparing the size-magnitude distribution of objects from our deeper surveys with our shallower surveys, it is obvious that significant incompleteness only sets in beyond $z_{850, AB} \sim 26.8$ in the GOODS fields and $z_{850, AB} \sim 27.5$ in the UDF-Ps. A more detailed quantification of these biases are provided in the text and Table D17 (see also Bouwens et al. 2004b).

### Table D15

<table>
<thead>
<tr>
<th>Magnitude Interval</th>
<th>GOODS$^b$</th>
<th>UDF-Ps$^b$</th>
<th>UDF</th>
</tr>
</thead>
<tbody>
<tr>
<td>$24.5 &lt; z_{850, AB} &lt; 25.0$</td>
<td>1.0</td>
<td>1.0</td>
<td>1</td>
</tr>
<tr>
<td>$25.0 &lt; z_{850, AB} &lt; 25.5$</td>
<td>0.0</td>
<td>0.0</td>
<td>0</td>
</tr>
<tr>
<td>$25.5 &lt; z_{850, AB} &lt; 26.0$</td>
<td>0.0</td>
<td>0.0</td>
<td>0</td>
</tr>
<tr>
<td>$26.0 &lt; z_{850, AB} &lt; 26.5$</td>
<td>2.7</td>
<td>2.9</td>
<td>3</td>
</tr>
<tr>
<td>$26.5 &lt; z_{850, AB} &lt; 27.0$</td>
<td>3.1</td>
<td>7.0</td>
<td>7</td>
</tr>
<tr>
<td>$27.0 &lt; z_{850, AB} &lt; 27.5$</td>
<td>1.8</td>
<td>7.9</td>
<td>8</td>
</tr>
<tr>
<td>$27.5 &lt; z_{850, AB} &lt; 28.0$</td>
<td>0.6</td>
<td>13.7</td>
<td>24</td>
</tr>
<tr>
<td>$28.0 &lt; z_{850, AB} &lt; 28.5$</td>
<td>0.0</td>
<td>4.9</td>
<td>16</td>
</tr>
<tr>
<td>$28.5 &lt; z_{850, AB} &lt; 29.0$</td>
<td>0.0</td>
<td>2.8</td>
<td>46</td>
</tr>
</tbody>
</table>

$^a$The figures here correspond to the number of $i$-dropout objects found in degradations of the UDF (Appendix C and Appendix D.1) to different depths (or as found in the original data). The degradation experiments were repeated 10 times, which is why the quoted values are often non-integer. This demonstrates how completeness can depend upon depth.

$^b$All of the dropouts listed in these columns were all identified as objects in our UDF catalogs. This ensures that differences in the deblending with foreground galaxies do not have a large effect on these results.

profiles) of these dropouts in the simulations to avoid introducing additional noise onto the image frames. To control for possible incompleteness from photometric scatter and surface brightness selection effects, we also laid down dropouts on empty frames. The net increase in incompleteness due to the presence of foreground objects is approximately 17%, 10%, and 8% for $i$-dropout objects in our UDF, UDF-P, and GOODS fields. These numbers appear to be relatively insensitive
to the flux of the source.\textsuperscript{6}

As a basic check on these results and to see how much our incompleteness determinations were affected by our choice of deblending parameters, we experimented with a smaller value for the deblending parameter (DEBLEND\_MINCONT=0.0001). With this choice, we calculated an incompleteness of 11\%, again using the above MINCONT, SExtractor rarely deblends sources. As a result, objects whose profiles overlap at or above some minimum flux threshold will be blended together. Since these thresholds occur at lower intrinsic surface brightnesses in our deeper fields, the blending will also be larger there.

\begin{table}
\centering
\caption{Number of i-dropouts from the UDF-Ps (20 arcmin\(^2\)) recovered at the depths of the GOODS fields.\textsuperscript{a}}
\begin{tabular}{llll}
\hline
Magnitude Interval & GOODS\textsuperscript{b} & UDF-Ps \\
\hline
24.5 < \textit{z}\textsubscript{850,\,AB} < 25.0 & 0.0 & 0 \\
25.0 < \textit{z}\textsubscript{850,\,AB} < 25.5 & 0.0 & 0 \\
25.5 < \textit{z}\textsubscript{850,\,AB} < 26.0 & 0.0 & 0 \\
26.0 < \textit{z}\textsubscript{850,\,AB} < 26.5 & 2.0 & 3 \\
26.5 < \textit{z}\textsubscript{850,\,AB} < 27.0 & 2.4 & 3 \\
27.0 < \textit{z}\textsubscript{850,\,AB} < 27.5 & 2.9 & 15 \\
27.5 < \textit{z}\textsubscript{850,\,AB} < 28.0 & 1.1 & 27 \\
28.0 < \textit{z}\textsubscript{850,\,AB} < 28.5 & 0.1 & 29 \\
28.5 < \textit{z}\textsubscript{850,\,AB} < 29.0 & 0.0 & 3 \\
\hline
\end{tabular}
\textsuperscript{a}The figures here correspond to the number of i-dropouts found in degradations of the UDF-Ps (Appendix C and Appendix D.1) to the depth of the two GOODS fields (or as found in the original data). The degradation experiments were repeated 10 times, which is why the quoted values are often non-integer. This demonstrates how completeness can depend upon depth.
\end{table}

\begin{table}
\centering
\caption{The Relative Completeness of the shallower data sets to the UDF.}
\begin{tabular}{llllll}
\hline
Magnitude Interval & GOODS (OBS)\textsuperscript{a} & GOODS (SIM)\textsuperscript{b} & UDF-Ps (OBS)\textsuperscript{a} & UDF-Ps (SIM)\textsuperscript{b} \\
\hline
24.5 < \textit{z}_{850,\,AB} < 25.0 & > 0.56 & 0.98 & > 0.56 & 1.00 \\
25.0 < \textit{z}_{850,\,AB} < 25.5 & 0.92 & 0.92 & 0.99 & 0.99 \\
25.5 < \textit{z}_{850,\,AB} < 26.0 & 0.86 & 0.86 & 0.86 & 0.86 \\
26.0 < \textit{z}_{850,\,AB} < 26.5 & 0.77^{+0.13}_{-0.17} & 0.79 & 0.97^{+0.03}_{-0.24} & 0.98 \\
26.5 < \textit{z}_{850,\,AB} < 27.0 & 0.55^{+0.14}_{-0.14} & 0.63 & > 0.87 & 0.90 \\
27.0 < \textit{z}_{850,\,AB} < 27.5 & 0.22^{+0.10}_{-0.07} & 0.32 & 0.99^{+0.01}_{-0.12} & 0.86 \\
27.5 < \textit{z}_{850,\,AB} < 28.0 & 0.04^{+0.04}_{-0.02} & 0.05 & 0.57^{+0.10}_{-0.10} & 0.61 \\
28.0 < \textit{z}_{850,\,AB} < 28.5 & < 0.03 & 0.00 & 0.31^{+0.11}_{-0.10} & 0.28 \\
28.5 < \textit{z}_{850,\,AB} < 29.0 & < 0.02 & 0.00 & 0.06^{+0.04}_{-0.03} & 0.04 \\
\hline
\end{tabular}
\textsuperscript{a}The relative completeness shown in column “OBS” relies upon the numbers obtained from the degraded data (Tables D15-D16). 1\sigma errors are calculated assuming binomial statistics (Appendix D.1). Lower limits are 1\sigma.
\textsuperscript{b}The relative completeness shown in column “SIM” are based upon the simulations we use to compute the transfer functions (Appendix D.3). Good agreement is observed relative to those extracted from the data, suggesting that the transfer functions we derive from these simulations are accurate.
\end{table}

\textsuperscript{6} Notice that the incompleteness is slightly larger for our deeper fields than for the shallower fields. This can attributed to our choice of deblending parameters (i.e., DEBLEND\_MINCONT=0.15). For such large values of DEBLEND\_MINCONT, SExtractor rarely deblends sources. As a result, objects whose profiles overlap at or above some minimum flux threshold will be blended together. Since these thresholds occur at lower intrinsic surface brightnesses in our deeper fields, the blending will also be larger there.
TABLE D18

i-dropouts in the UDF which were blended with foreground galaxies in our main UDF catalog (Table 3). *

<table>
<thead>
<tr>
<th>Object ID</th>
<th>RA</th>
<th>Dec</th>
<th>z_{5850}</th>
<th>i - z</th>
<th>z - J</th>
<th>J - H</th>
<th>S/G</th>
<th>r_{hi}(′′)</th>
</tr>
</thead>
<tbody>
<tr>
<td>UDF-38397588</td>
<td>03:32:38.39</td>
<td>-27:47:58.8</td>
<td>26.94±0.04</td>
<td>1.3</td>
<td>-0.2</td>
<td>0.0</td>
<td>0.02</td>
<td>0.14</td>
</tr>
<tr>
<td>UDF-36456342</td>
<td>03:32:36.45</td>
<td>-27:48:34.2</td>
<td>27.25±0.06</td>
<td>1.3</td>
<td>&lt;0.1 faint</td>
<td>0.01</td>
<td>0.17</td>
<td></td>
</tr>
<tr>
<td>UDF-3556441</td>
<td>03:32:33.55</td>
<td>-27:46:44.1</td>
<td>27.27±0.07</td>
<td>1.3</td>
<td>0.0</td>
<td>0.00</td>
<td>0.19</td>
<td></td>
</tr>
<tr>
<td>UDF-3727545</td>
<td>03:32:37.27</td>
<td>-27:48:54.5</td>
<td>27.48±0.05</td>
<td>3.0</td>
<td>—</td>
<td>—</td>
<td>0.22</td>
<td>0.12</td>
</tr>
<tr>
<td>UDF-42543938</td>
<td>03:32:42.54</td>
<td>-27:48:39.8</td>
<td>27.74±0.08</td>
<td>1.5</td>
<td>—</td>
<td>—</td>
<td>0.01</td>
<td>0.15</td>
</tr>
<tr>
<td>UDF-33556440</td>
<td>03:32:33.55</td>
<td>-27:46:44.0</td>
<td>27.86±0.08</td>
<td>2.2</td>
<td>—</td>
<td>—</td>
<td>0.01</td>
<td>0.14</td>
</tr>
</tbody>
</table>

*Table D18 is published in its entirety in the electronic version of the Astrophysical Journal. A portion is shown here for guidance regarding its form and content. Similar comments to Table 3 apply. Sources to our main catalogs were found (Appendix D.1) using a more aggressive splitting parameter (DEBLEND_MINCONT=0.0001) than used in the main catalog (DEBLEND_MINCONT=0.15; see §3). Adding these sources to our main catalogs would increase the total number of i-dropouts in the HUDF by ~ 7%.

D.2. Flux Corrections

Depth can also have an impact on the measured magnitudes. This is particularly true for scalable aperture magnitudes (MAG_AUTO) as used by SExtractor, where both the shape and size of the aperture is set by the light above some isophote. Fainter lower surface brightness objects tend to have significantly smaller isophotal areas, and this can bias the size of the aperture derived to measure fluxes. To estimate the extent of this bias, we compared the z_{5850}-band magnitudes measured for specific i-dropouts in the UDF with measurements made on the same objects degraded down to GOODS and UDF-Ps depths and plotted these differences as a function of magnitude (Figure D20). Again, we considered the results of 10 different degradation experiments in constructing this plot (see Appendix C or Appendix D.1 for a description). Despite considerable amounts of scatter, magnitudes measured in the UDF were found to be ~ 0.1 mag and ~ 0.2 mag brighter than that measured at UDF-Ps and GOODS depth. Near the selection limit, there was a noticeable decrease in the mean flux bias. This appears to be the result of a Malmquist-like selection effect (i.e., near the magnitude limit, bright scattering objects make it into our selection while faintward scattering objects do not). We compiled the results of these experiments into an average offset vs. magnitude (red vertical bars in Figure D20). The 68% confidence limits on these offsets were derived from the object-to-object scatter.

D.3. Transfer Functions

The completeness and flux corrections detailed in Appendix D.1 and D.2 can be more properly implemented using transfer functions. Transfer functions take surface densities observed at one depth and convert them to their equivalent densities if measured at another. In this formulation, incompleteness is incorporated as a decrease in the surface density from the input to output stage. Magnitude biases are included by effecting a shift from one magnitude interval to another.

Ideally, we would determine the transfer functions in the same way as we estimated the completeness and flux biases in the previous sections (e.g., by performing degradation experiments on the real data). Unfortunately, the available data are simply not sufficient to adequately determine these functions. Without a large number of input objects, the computed transfer functions would be overly dependent on the position of particular objects within the different magnitude bins (and their morphologies), compromising the overall accuracy of the simulations. This is particularly true at bright magnitudes (z_{5850,AB} < 26) where there is only one object in our deeper fields.

As such, it appeared that our best option was simply to rely on simulations—again using our cloning machinery to generate the mock fields. The inputs to the simulations consisted of B-dropout samples from both the UDF-Ps (Bouwens et al. 2004b) and the HUDF (Bouwens et al. 2004b). Our use of z ~ 3.8 B-dropout samples was motivated by the much higher surface brightness sensitivities available for B-dropouts than for i-dropouts in the same data (due to (1 + z)^4 cosmic surface brightness dimming). Moreover, objects from these samples should be fairly similar to the i-dropout sample both in size and morphology, minimizing the important of different assumptions regarding their evolution over cosmic time (~700 Myr). Objects were projected over the range z ~ 5.2 – 7.0 in accordance to their volume density and then added to the artificial UDF frames. Object sizes were scaled as (1 + z)^{-1.1} (for fixed luminosity) to match the observed scalings (§3.7).

Our transfer functions were calculated by degrading the above simulations and then comparing the magnitudes of objects selected on the original frames (at UDF depths) with those selected on the degraded frames. The transfer functions are initially binned on 0.1 magnitude scales to form familiar two-dimensional matrices, and then smoothed along the diagonals (to improve the statistics while preserving flux biases). The smoothing length is set so that at least 30 different objects from our simulated images contribute to each element in this matrix (this is equivalent to Δm = 0.5 at z_{5850,AB} < 25.5, but Δm = 0.1 at z_{5850,AB} > 26.5). After smoothing, the results are rebinned in 0.5 magnitude intervals to match the binning for the number counts (Figure 6). A tabulation of our two transfer functions are provided in Tables D19 and D20. They are expressed in such a way that one can use matrix multiplication procedures to go from surface densities selected in UDF-type data to the equivalent surface densities measured in the

<table>
<thead>
<tr>
<th>Object ID</th>
<th>Dec</th>
<th>RA</th>
<th>z_{5850}</th>
<th>i - z</th>
<th>z - J</th>
<th>J - H</th>
<th>S/G</th>
<th>r_{hi}(′′)</th>
</tr>
</thead>
<tbody>
<tr>
<td>UDF-38397588</td>
<td>-27:47:58.8</td>
<td>03:32:38.39</td>
<td>26.94±0.04</td>
<td>1.3</td>
<td>-0.2</td>
<td>0.0</td>
<td>0.02</td>
<td>0.14</td>
</tr>
<tr>
<td>UDF-36456342</td>
<td>-27:48:34.2</td>
<td>03:32:36.45</td>
<td>27.25±0.06</td>
<td>1.3</td>
<td>&lt;0.1 faint</td>
<td>0.01</td>
<td>0.17</td>
<td></td>
</tr>
<tr>
<td>UDF-3556441</td>
<td>-27:46:44.1</td>
<td>03:32:33.55</td>
<td>27.27±0.07</td>
<td>1.3</td>
<td>0.0</td>
<td>0.00</td>
<td>0.19</td>
<td></td>
</tr>
<tr>
<td>UDF-3727545</td>
<td>-27:48:54.5</td>
<td>03:32:37.27</td>
<td>27.48±0.05</td>
<td>3.0</td>
<td>—</td>
<td>—</td>
<td>0.22</td>
<td>0.12</td>
</tr>
<tr>
<td>UDF-42543938</td>
<td>-27:48:39.8</td>
<td>03:32:42.54</td>
<td>27.74±0.08</td>
<td>1.5</td>
<td>—</td>
<td>—</td>
<td>0.01</td>
<td>0.15</td>
</tr>
<tr>
<td>UDF-33556440</td>
<td>-27:46:44.0</td>
<td>03:32:33.55</td>
<td>27.86±0.08</td>
<td>2.2</td>
<td>—</td>
<td>—</td>
<td>0.01</td>
<td>0.14</td>
</tr>
</tbody>
</table>
shallower data. Note that since object blending is not properly included in these simulations (the surface density of objects is comparably low), we corrected our transfer functions up to account for the greater incompleteness in the UDF due to this blending (see Appendix D.1).

It is possible to obtain a useful check on the results obtained from these simulated fields by estimating the completeness levels and flux biases on these same fields. Our estimates of these quantities were computed in a very similar way as on the actual data (i.e., Appendix D.1 and Appendix D.2) to ensure consistency. The results are shown in column “SIM” of Table D17 and Figure D20 (blue shaded regions) and appear to be in broad agreement with those obtained from our degradation experiments. This provides us with confidence in the transfer functions we determine from the simulations.

D.4. Contamination Corrections

In principle, the availability of $B_{435}$ and $V_{606}$-band imaging provides an effective means of eliminating lower redshift contaminants directly. Lower redshift interlopers are expected to be significantly brighter in the $B_{435}$ and $V_{606}$-bands than genuine high-redshift objects and therefore our requirement that objects be redder than 2.8 in $V_{606} - z_{850}$ (Figure A18, Appendix A) should prove to be an effective means of eliminating such objects. Unfortunately, near the magnitude limit of each field, only limited constraints can be set on the $V_{606}$-band fluxes and therefore it is difficult to effectively filter out all contaminants.

Though there is some indication that the flux biases we derive from the simulations may underestimate those obtained from the observations (top panel of Figure D20), this may simply be an artefact of the objects we use to make these estimates (only 3 objects from the HUDF were used to derive the mean flux biases in the brightest two magnitude bins). Since possible systematics are much smaller in size than the uncertainties due to large-scale structure (i.e., $\sigma(M_{1350}) \approx 0.15$: Appendix E), we ignore this issue for the rest of this analysis.
We can however estimate this contamination statistically, using the deeper optical and infrared data available for some of our fields. We break these contamination estimates into four different components: (1) contamination from intrinsically red objects, (2) contamination from photometric scatter, (3) contamination from lower mass stars, and (4) contamination from spurious sources. Explicit effort is made to ensure that the contribution from each component is independent (and thus no contaminant is subtracted twice).

D.4.1 Contamination from Intrinsically Red Objects

A small fraction of low-redshift \((z \sim 1 - 3)\) galaxies have colors which are red enough to satisfy our \((i_{775} - z_{850})_{AB} > 1.3\) selection. Since such objects will also have red \(z - K_s\) colors (while bona-fide \(z > 5.5\) objects are likely to be much bluer), we can estimate the contamination rate from these objects by examining these colors for a sample of \(i\)-dropout candidates. As shown in Figure 3 and D21, the optical-infrared colors provide a good separation between high-redshift objects and lower redshift interlopers. Though we already provided a preliminary estimate of this contamination rate from the HUDF in §3.2, we can obtain a much better estimate of this contamination rate at bright magnitudes \((25 < z_{850,AB} < 27)\) using the ISAAC data available over the CDF-S GOODS field. Similar to the procedures outlined at the beginning of §3, \(z_{850} - J\) and \(z_{850} - K_s\) colors for \(i\)-dropouts in CDF-South GOODS were measured by smoothing the \(z_{850}\)-band data to the same PSF as in the infrared and measuring the flux in an aperture whose diameter was \(2\times\) the FWHM of the object. Compiling galaxies from the entire 131 arcmin\(^2\) CDF-S ISAAC mosaic, candidate low-\(z\) interlopers were identified with a \((i_{775} - z_{850})_{AB} > 1.3; (z_{850} - K_s)_{AB} > 1.6\) criterion. Only two such objects were found (Figure D21): one at \(z_{850,AB} \sim 25.4\) and one at \(z_{850,AB} \sim 26.0\). The majority of objects with \(i_{775} - z_{850} > 1.3\) colors had \((z_{850} - K_s)_{AB}\) colors less than 1.6. Over the interval \(25.0 < z_{850,AB} < 26.0\), this works out to \(18^{+13}_{-9}\%\) contamination rate from intrinsically red objects and over the interval \(26.0 < z_{850,AB} < 27.0\), the contamination rate is \(\lesssim 2\%\) (1\(\sigma\)). These results are combined with similar estimates from the UDF IR data (§3.2) and summarized in Table D21.
Fig. D21.— \((i_{775} - z_{850})_{AB}/(z_{850} - K_s)_{AB}\) colors of objects in the CDF-South GOODS field (Appendix D.4.1) with \(z_{850,AB} < 26.8\). Objects which are undetected (<2σ) in the \(V_{606}\)-band are shown in black while objects which are detected at the 2σ level are shown in magenta. Objects which made it into our low-redshift interloper selection (Table C14: \((i_{775} - z_{850})_{AB} > 1\) and \((z_{850} - J)_{AB} > 0.8\)) are shown as enlarged open squares (see also Figure A18). Otherwise objects are shown as small filled circles. Color-color tracks of low-redshift templates and high-redshift starbursts with different reddenings are as in Figure 3. Arrows denote 2σ limits on the \((i_{775} - z_{850})_{AB}\) and \((z_{850} - K_s)_{AB}\) colors. The solid horizontal line shows our \((i_{775} - z_{850})_{AB}\) cut for selecting \(i\)-dropouts while the solid vertical line shows our \((z_{850} - K_s)_{AB}\) cut which serves to separate dropouts from intrinsically red objects (Figure A18). The majority (<2%) of objects with \(i_{775} - z_{850} \geq 1.3\) colors had \((z_{850} - K_s)_{AB}\) colors less than 1.6. This suggests that contamination from intrinsically red objects is very small (<2%; Table D21).

### Table D21

<table>
<thead>
<tr>
<th>Magnitude Interval</th>
<th>GOODS(^a)</th>
<th>UDF(^a)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(24.0 &lt; z_{850,AB} &lt; 25.0)</td>
<td>0.000 (&lt;43%)</td>
<td>0.000 (&lt;43%)</td>
</tr>
<tr>
<td>(25.0 &lt; z_{850,AB} &lt; 26.0)</td>
<td>0.012 (18(^{+13}_{-9})%</td>
<td>–</td>
</tr>
<tr>
<td>(26.0 &lt; z_{850,AB} &lt; 27.0)</td>
<td>0.000 (&lt;2%)</td>
<td>0.000 (&lt;17%)</td>
</tr>
<tr>
<td>(27.0 &lt; z_{850,AB} &lt; 28.0)</td>
<td>–</td>
<td>0.000 (&lt;6%)</td>
</tr>
<tr>
<td>(28.0 &lt; z_{850,AB} &lt; 29.0)</td>
<td>–</td>
<td>0.050 (10(^{+8}_{-5})%</td>
</tr>
<tr>
<td>(29.0 &lt; z_{850,AB} &lt; 29.5)</td>
<td>–</td>
<td>–</td>
</tr>
</tbody>
</table>

\(^a\) Units are arcmin\(^{-2}\). The number in parenthesis indicates the fraction of \(i\)-dropout candidates with optical-infrared colors suggesting that they are intrinsically red low redshift contaminants (Appendix D.4.1, §3.2). Uncertainties are 1σ and were determined from binomial statistics.

### D.4.2. Contamination from Photometric Scatter

Here we estimate the contamination rate from photometric scatter. This accounts for objects whose measured colors satisfy our selection criteria in both the optical and the infrared, but for which deeper observations would reveal them to be at lower redshift (i.e., have colors more consistent with a lower redshift object). As with our estimates of the completeness levels and flux biases, perhaps the most model-independent procedure is to use the results of our degradation experiments described earlier (Appendix C). Objects which are selected as \(i\)-dropouts can be compared against the original source catalogs available for the UDF and UDF-Ps fields and contaminants identified. Note that objects are only classified as contaminants if the deeper photometry suggests that their redshifts are likely well below 5.0, i.e., significantly below our nominal lower redshift limit of \(z \sim 5.5\) (Figure 10). This will happen for \((V_{606} - z_{850})_{AB}\) colors bluer than 2.8 (Appendix A). Having \((i_{775} - z_{850})_{AB}\) colors bluer than 1.3 (in the deeper photometry) is not
sufficient to label an object a contaminant. This avoids classifying as contaminants objects which are just below our nominal low redshift limit (\(z \approx 5\)); see Figure A18) and thus readily scatter into our selection. The results of these simulations are compiled in Tables D22-D23 as a function of magnitude, and again this source of contamination is small (\(\lesssim 10\%\)) and only of significance within \(\pm 1\) mag of the faint-end limit.

In our shallower fields, contamination from photometric scatter can effectively be controlled for using degradations of the UDF. However, the UDF itself has no deeper field which can serve as a control (particularly faintward of \(z_{850,AB} \sim 28.5\) where the UDF \(V_{606}\)-band fluxes are no longer of sufficient S/N to filter out contaminants). Therefore, we needed an alternative procedure, and so we elected to model the faint objects in our catalog with the colors of intermediate magnitude \(25.9 < z_{850,AB} < 27.4\) objects and then add photometric scatter. To ensure that the intermediate magnitude objects were really at low redshift, we required the objects to have \((i_{775} - z_{850})_{AB}\) colors bluer than 0.9 and \((V_{606} - z_{850})_{AB}\) colors bluer than 2.5. These criteria explicitly excluded objects which were close to qualifying as \(i_{775}\)-dropouts (see Figure D22). In performing the simulations, we iterated over all faint \(z_{850,AB} > 27.9\) objects in the UDF (2908 objects), randomly picking an intermediate magnitude object and then perturbing this object’s photometry to match the S/N of the faint object we were iterating over. After repeating this scattering experiment on all faint objects in the UDF in four separate trials, we found only one contaminant, or just 0.25 contaminant per 13 arcmin\(^2\) field. This is a smaller fraction than what we found in our simulations of the UDF-Ps and GOODS fields (Tables D22-D23) and may owe to the depth of the UDF \(i_{775}\)-band imaging. In our other fields, the \(i_{775}\)-band depths only exceeded the \(z_{850}\)-band depths by \(\sim 0.4\) mags, but in the UDF this difference is 0.8 mags. Also note that because at faint magnitudes almost all objects are blue \((i_{775} - z_{850}) \lesssim 0.6\), \((V_{606} - z_{850})_{AB} < 1.3\), most objects would still be quite significant detections in the bluer bands at the limits of the UDF \(i_{775}\)-dropout probe (\(z_{850,AB} \sim 29.5\)).
Low mass stars have similar $(i' - z_{850})_{AB}$ colors to $z \sim 6$ objects, and therefore can act as contaminants to our samples. Fortunately, this has not proven to be a large concern, mostly because the majority of $i$-dropouts ($\gtrsim 90\%$) are clearly resolved at ACS resolution ($0.10''$ FWHM) and therefore it has been straightforward to separate these objects (which have typical half-light radii of $\sim 0.1 - 0.2''$) from stellar contaminants. This separation is particularly clear at bright magnitudes or high S/N where even for marginally resolved objects it is possible to measure some extension or diffuse emission around the object. At fainter magnitudes, however, the situation becomes a little more challenging, especially in our shallower surveys or at lower S/N.

In these cases, we think it is best to use our deeper fields as a guide to the true contamination rate and not attempt to derive it from the shallower data. So instead of using the measured stellarities to decide whether to include or exclude objects from our catalogs, we simply estimate a contamination fraction from the deeper data and apply that to the surface densities measured from our shallower data. Again this statistical approach is most useful at magnitudes where the signal-to-noise no longer allows a clear separation between point-like and extended objects, i.e., $z_{850,AB} \gtrsim 26.8$ for the GOODS fields, $z_{850,AB} \gtrsim 27.5$ for the UDF-Ps, and $z_{850,AB} \gtrsim 28.4$ for the UDF. An estimate of this contamination fraction can be obtained by examining the data at all three depths and plotting the fraction of point-like objects as a function of the $z_{850}$-band magnitude in all three fields (Figure D23). Here there is a clear, near monotonic decrease in the fraction of these objects with magnitude, from $\sim 80 - 100\%$ at bright magnitudes ($z_{850,AB} \sim 23 - 25$) to a mere $\sim 1-2\%$ at fainter magnitudes ($z_{850,AB} \sim 26 - 27$).

One exception to this trend occurs near the selection limit of each sample, where the pointlike fraction is observed to rise. While this must be partially due to a selection bias for small objects near the magnitude limit (see Figure D19), it is also one likely result of the increasingly poor measure one has of the stellarity at these S/N. Are these objects stars? A useful way of checking this is to look at their optical-infrared colors shown in Figure D24. If these objects were stars, the $z - J$ colors would be similar to other low mass stars observed in the GOODS fields (where the $z - J$ colors are $\gtrsim 1$ mag and more likely $\gtrsim 1.5$ mag to match the expected color of lower luminosity, lower mass T dwarfs more common at faint magnitudes). As it turns out, none of the four point-like objects at $26.8 < z_{850,AB} < 27.4$ has

---

**Figure D22.** (left panel) The intermediate magnitude photometric sample (gray shaded region in the lower left hand corner) used to estimate the susceptibility of our faint ($z_{850,AB} > 27.9$) UDF $i$-dropout sample to contamination from low-redshift interlopers (due to photometric scatter). Shown are the $(i' - z_{850})_{AB}/(V_{606} - z_{850})_{AB}$ colors of faint $25.9 < z_{850,AB} < 27.4$ objects from the UDF (black circles). The horizontal and vertical lines show the $(i' - z_{850})_{AB}$ and $(V_{606} - z_{850})_{AB}$ selection cuts used for selecting $i$-dropouts. Objects that are particularly red ($\gtrsim 0.7$) in $(z_{850} - J_{110})_{AB}$ and therefore likely low-redshift early-types are shown as red circles. $2\sigma$ lower limits are indicated with the arrows. The color-color tracks for low-redshift interlopers are also included along with the position of high-redshift starbursts with various amounts of reddenings. Objects in the upper right hand corner (orange shaded region) are $i$-dropouts, and objects in the lower-right hand corner are objects which are likely just below our redshift cut. The position of two $i$-dropouts that are partially blended with foreground objects are indicated by the blue arrow (Table D18) while the position of one point-like star is indicated by the red cross. (right panel) $(i' - z_{850})_{AB}/(V_{606} - z_{850})_{AB}$ colors for objects from the HUDF input sample (lower lefthand corner with gray shading) scattered to match the photometric errors of faint ($z_{850,AB} > 27.9$) objects in the HUDF. The output of the simulations indicates that contamination from photometric scatter at faint magnitudes is negligible ($< 1$ object) (see Appendix D.4.2).
Fig. D23.— Fraction of \(i\)-dropout candidates that are point-like (SExtractor stellarity > 0.75, where 0 = an extended object and 1 = a point source) and thus likely stellar contaminants vs. \(z_{850}\)-band magnitude. (Left panel) Observed fraction in the UDF (red lines), UDF-Ps (blue lines), and GOODS fields (black lines). The lines become dotted at the point where the S/N is too low to discriminate between extended and point-like objects. (Inset) The assumed fraction of stellar contaminants (shaded red region shows the assumed 1σ uncertainties). Stellar contaminants are rejected using the measured stellarties brightward of \(z_{850},AB\) equal to 28.4, 27.5, and 26.8, for the UDF, UDF-Ps, and GOODS fields, respectively. Faintward of this, no such attempt is made and a contamination fraction is assumed based upon an extrapolation from bright magnitudes (see Appendix D.4.3 for more details).

a detection in the \(J\)-band and most have \((z_{850} − J)_{AB}\) colors with 2σ limits blueward of 0.5. This suggests that most of the fainter objects in the GOODS fields are extended despite their rather pointlike appearance and therefore the contamination rate from low-mass stars remains low as indicated by our deeper ACS data. Similar conclusions hold for faint objects in the UDF (\(z_{850,AB} < 29\): right panel to Figure D24).

Connecting the surveys up and extrapolating the trends beyond \(z_{850,AB} ∼ 28.4\), we can arrive at an approximate contamination fraction as a function of magnitude (right panel to Figure D23). Multiplying these fractions by the observed surface densities (Table 10), the contamination rate from low mass stars can be derived.

D.4.4. Contamination from Spurious Sources

In principle, our samples were also sensitive to contamination from spurious objects resulting from noise spikes or other non-gaussian features. If present, such spurious sources would easily qualify as dropouts given the unlikelihood that similar spikes would occur in the other passbands. Therefore, analogous to the simulations described in B05b, Dickinson et al. (2004), and Yan & Windhorst (2004b), we repeated our selection procedure on the negative images, and similar to the results in B05b and Yan & Windhorst (2004b), no objects were found in our data sets at all three depths. Therefore, it seems unlikely that spurious objects represent a significant source of contamination for our samples (\(< 1\%)\).

D.4.5. Summary

Table D24 shows the sum of all three sources of contamination for the samples considered here (spurious sources do not appear to be a concern). Totalling up these results for all 3 samples and all magnitude intervals, we can arrive at an approximate contamination rate for our cumulative sample. This result is \(< 8\%\) (i.e., \(> 92\%\) of \(i\)-dropouts are at \(z ∼ 6\)).

E. UNCERTAINTIES IN THE LUMINOSITY FUNCTION DUE TO FIELD-TO-FIELD VARIATIONS

In deriving the rest-frame continuum UV luminosity function at \(z ∼ 6\), we make use of \(i\)-dropouts from three different fields. One possibly significant concern is that since the surface density of \(i\)-dropouts can show significant differences in normalization from one field to another (we expect ~35% RMS for a 12 arcmin$^2$ ACS field: §3.6), these differences may have an effect on our derived LF. To quantify the size of this effect, we ran a series of Monte-Carlo simulations. Using the normalization \(\phi^*_* = 0.00176\) Mpc$^{-3}$ and faint-end slope \(α = −1.74\) from our best-fit LF (§5) and an ensemble of different characteristic luminosities \(M^*_{1350,AB}\) (i.e., \(−19.75, −19.85, −19.95, ..., −20.65\)) scattered around our preferred value of \(M^*_{1350,AB} = −20.20\) (§5), we generated number count predictions for each of our fields (i.e., the UDF, the UDF-Ps, and the GOODS fields). Our computed counts included the relevant selection and measurement biases as shown in Tables D19-D20 and Figure 9. We varied the normalization on our counts for our deepest two fields (i.e., the UDF and UDF-Ps) by 30% RMS (the approximate uncertainties on the relative normalization of our different
Fig. D24.— (Left panel) \((z_{850} - J)_{AB}/(J - K_s)_{AB}\) colors of point-like objects (SExtractor stellarity > 0.75, where 0 = an extended object and 1 = a point source) in the CDF South GOODS field which satisfy our \((i_{775} - z_{850})_{AB} > 1.3\) cut. The solid black points show the position of bright objects \((z_{850,AB} < 24.9)\), the open black circles show the position of intermediate magnitude objects \((24.9 < z_{850,AB} < 26.8)\), and the solid red points show the position of fainter \((26.8 < z_{850,AB} < 27.4)\) objects. The shaded light green region shows the region where M, L, and T stars are expected to lie while the blue tracks show the colors of starbursts with various amounts of reddening. The colors of most faint \(850 < z_{850,AB} < 27.4\) point-like objects in GOODS (solid red points) are inconsistent with being low-mass stars. These results are all the more significant since most intermediate magnitude to faint objects (open circles) have \(z_{850} - J\) colors more consistent with lower mass T dwarfs, i.e., \((z_{850} - J)_{AB} \gtrsim 1.5\) mag. This suggests that contamination by stars remains low faintward of \(z \sim 26.8\) as indicated by our deeper fields (red and blue lines in left-hand panel of Figure D23). (Right panel) \((z_{850} - J_{110})_{AB}/(J_{110} - H_{160})_{AB}\) colors for point-like objects in the UDF which satisfy our \((i_{775} - z_{850})_{AB} > 1.3\) cut. The red points show the position of faint \(27.9 < z_{850,AB} < 29.2\) objects. Otherwise as in the left panel. Similar to conclusions drawn from the left panel, the \((z_{850} - J_{110})_{AB}\) colors observed for faint pointlike objects in the UDF \((z_{850,AB} < 29)\) are inconsistent with most of these objects being stellar in origin (see Appendix D.4.3).

<table>
<thead>
<tr>
<th>Mag. Range</th>
<th>GOODS</th>
<th>Field</th>
</tr>
</thead>
<tbody>
<tr>
<td>GOODSb</td>
<td></td>
<td>UDF-Psb</td>
</tr>
<tr>
<td>24.0 &lt; (z_{850}) &lt; 24.5</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>24.5 &lt; (z_{850}) &lt; 25.0</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>25.0 &lt; (z_{850}) &lt; 25.5</td>
<td>0.003</td>
<td>0.000</td>
</tr>
<tr>
<td>25.5 &lt; (z_{850}) &lt; 26.0</td>
<td>0.009</td>
<td>0.000</td>
</tr>
<tr>
<td>26.0 &lt; (z_{850}) &lt; 26.5</td>
<td>0.013</td>
<td>0.000</td>
</tr>
<tr>
<td>26.5 &lt; (z_{850}) &lt; 27.0</td>
<td>0.009</td>
<td>0.000</td>
</tr>
<tr>
<td>27.0 &lt; (z_{850}) &lt; 27.5</td>
<td>0.019</td>
<td>0.008</td>
</tr>
<tr>
<td>27.5 &lt; (z_{850}) &lt; 28.0</td>
<td>0.014</td>
<td>0.170</td>
</tr>
<tr>
<td>28.0 &lt; (z_{850}) &lt; 28.5</td>
<td>–</td>
<td>0.266</td>
</tr>
<tr>
<td>28.5 &lt; (z_{850}) &lt; 29.0</td>
<td>–</td>
<td>0.081</td>
</tr>
<tr>
<td>29.0 &lt; (z_{850}) &lt; 29.5</td>
<td>–</td>
<td>–</td>
</tr>
</tbody>
</table>

aSince the brighter stars \((z_{850} < 26.8, 27.5, \text{and} 28.4)\) for the GOODS fields, UDF-Ps, and UDF, respectively) are explicitly filtered out using the measured stellarieties (§3.1: Table 2), we assume no contribution to the contamination rate from stellar objects at these magnitudes. Faintward of these limits, the stellar contamination is assumed to be a declining fraction of the total surface density (Appendix D.4.3, Figure D23). 

bUnits are arcmin \(^{-2}\).

fields), combined the counts from all our fields (§3.8), and then fit them to a Schechter function (§5). Repeating this experiment 100 times using different normalizations for our three fields, we derived RMS errors on our three Schechter parameters that result from the uncertain normalizations. The RMS errors on \(\alpha\) were consistently \(~0.20\) for all input values of \(M_{1350}^*\), while the RMS errors on \(M_{1350}^*\) and \(\phi^*\) increased from 0.10 and 0.00035, respectively, for fainter values
of $M_{1350}^*$ (i.e., $-19.65$) to 0.17 and 0.00055, respectively, for brighter values of $M_{1350}^*$ (i.e., $-20.65$). This suggests that it is currently not possible to determine the normalization of the luminosity function $\phi^*$ to better than 30% and the faint-end slope $\alpha$ to better than 0.2.