PUBLICATIONS OF THE ASTRONOMICAL SOCIETY OF THE PACIFIC, **120**:275–280, 2008 February
© 2008. The Astronomical Society of the Pacific. All rights reserved. Printed in U.S.A.

QSOs in the Combined SDSS/*GALEX* Database

J. B. HUTCHINGS

Herzberg Institute of Astrophysics, Victoria, BC, Canada; john.hutchings@nrc.ca

AND

L. BIANCHI

Department of Physics and Astronomy, Johns Hopkins University, Baltimore, MD

Received 2007 November 02; accepted 2008 January 02; published 2008 February 26

ABSTRACT. We discuss selection of QSO candidates from the combined SDSS and *GALEX* catalogs. We discuss properties of QSOs within the combined sample and note uncertainties in number counts and completeness, compared with other SDSS-based samples. We discuss color and other properties with redshift within the sample and the spectral energy distributions for subsets. We estimate the numbers of faint QSOs that are classified as extended objects in the SDSS imaging and consequent uncertainties that follow.

1. INTRODUCTION

The Sloan Digital Sky Survey (SDSS) has produced major catalogs of optically selected objects that are of interest for many studies. The SDSS has also produced the largest collection of QSOs to date, by a large margin. The next step is to identify and verify which they are. While some of these objects have SDSS spectra that can identify them as QSOs, they are a small fraction compared with the *GALEX*-SDSS QSO candidate selection of Bianchi et al. (2007), and it is of interest to estimate the number counts with magnitude of QSOs, based on this larger and fainter data set. It is also interesting to examine the parameter space favored by the different selection criteria.

Like all optically selected QSO surveys, there are selection effects that must be understood, in order to produce these distributions correctly. The SDSS filters alone do not allow a clean separation of low-redshift QSOs from other blue objects. The *GALEX* UV surveys add very useful wavelength leverage to isolating QSO candidates, when combined with optical data (see Figs. 5–7 of Bianchi et al 2007). In this paper, we discuss QSO candidates based on the combined SDSS and *GALEX* survey databases. Specifically, we use the catalog of low-redshift QSO candidates selected from a matched source catalog constructed by Bianchi et al. (2007) from the *GALEX* GR1 and SDSS DR3 data releases. The overlap areas are 363 deg^2 in the *GALEX* All-sky Imaging Survey (AIS) and 83 deg^2 in the *GALEX* Medium Imaging Survey (MIS). The AIS has typical exposures of about 100 s and reaches objects of limiting (5 σ) NUV-band AB magnitudes of about 20.8 while the MIS has average exposures of 1500 s, reaching objects with NUV magnitudes of about 22.7.

Based on initial color cuts in FUV $-$ NUV and NUV $- r$ as in Bianchi et al (2007), and photometric error cuts of 0.3 mag in FUV, NUV, and r, we begin with about 34,000 objects from the MIS and 6000 from the AIS. Of these, 17% and 55%, respectively, are classified by SDSS as point sources. A further cut in u magnitude at 24.2 (or u-band error 1.5 mag) could eliminate a few more sources that are probably spurious in that band, but this amounts to less than 4% of the MIS point sources and none at all of the other candidate samples. In the sections below, we describe further cuts to these samples, to match the properties of the subsets of about 800 in the MIS and 1600 in the AIS that are confirmed QSOs, with redshifts, from SDSS spectra. We then discuss the properties of the resulting candidate catalogs in terms of the number counts of QSOs over the redshift range about 0.1 to 2. Table 1 gives a summary of the catalogs we discuss.

The Bianchi et al (2007) selection of QSOs should include essentially all QSOs between $z = 0.5$ and 1.5 (which are also the least contaminated by foreground stars), but it is definitely incomplete for z around 0 and >1.6, because the natural spread of spectral energy distributions (SEDs) around an average template, and extinction effects, blur QSO and stellar loci together. In other words, QSOs around $z = 0$ and >1.5 share the stellar locus in part, so their color selection was made so as to minimize the stellar contamination for $z = 0$, at the expense of missing a fraction of QSOs (see also Fig. 5 of Bianchi 2008). In this paper we discuss ways to get more complete QSO candidate lists and some values for their magnitude and redshift distribution.

2. PROPERTIES OF KNOWN QSOS

We recall that our sample of QSO candidates was selected by Bianchi et al. (2007) from the NUV $- r$ and FUV $-$ NUV color-color diagram (see their Fig. 7, *bottom panels*), where the colors of QSOs based on templates from previous known QSO samples separate from most types of stellar objects and galaxies. Therefore, this work addresses the selection of QSOs with

TABLE 1

GALEX-SDSS CANDIDATES AND CONFIRMED—DATABASE SIZES

Property	MIS ext	MIS point	AIS ext	AIS point
Total candidates	29287	5960	2633	3257
Sky area (sq deg)	83		363	
Spectra in total	0.2%	14%	16%	47%
QSO spectra[a]	34	813	109	1497
% QSO spectra	45%	85%	25%	81%
Galaxy spectra	42	8	317	21
Star spectra	0	132	8	335
QSO z < 0.7	30	154	103	493
QSO z 0.9 to 1.6[b]	...	412	...	623
QSO $z > 1.44$[c]	...	213	...	201
QSOs, color cuts				
$g-r$ full range[a]	21	762	58	1438
$g-r$ plus UV-opt[b]	12	382	38	508
$r-i$ plus UV-opt[c]	21	235	57	180
Selected candidates				
$r-i$, UV-opt	(10554)	1120	(1575)	230
$g-r$, UV-opt	(4945)	1646	(1188)	661

NOTE.—Matching superscripts a, b, and c indicate numbers to be compared. Numbers in parentheses are from extended sources with same cuts.

similar properties. In future work we will investigate the possible existence and characteristics of QSO samples with differing SEDs in the UV.

We use the photometric properties of the spectroscopically confirmed QSOs included in our candidate QSO catalogs to estimate the QSO population and properties of the full GALEX + SDSS catalogs. Figure 1 shows some of the photometric properties of the spectroscopically confirmed QSOs with redshift.

At redshifts 0.5 and higher, there is a very tight and systematic change of $g-r$ color with redshift. Lower redshift QSOs show scatter to larger values, but a $g-r$ range of -0.2 to 0.4 includes over 97% of them. There is a similar tight relationship with $r-i$. Figure 1 shows these samples.

The lower panel of Figure 1 shows the correlation involving UV flux and redshift. The quantity plotted is the following combination of magnitudes: $4r+2g-2i-z-3$FUV. This was found by trial to have a tighter and more single-valued dependence on redshift than any simple difference of two magnitudes, such as FUV $-$ u or NUV $-$ r, etc. However, these other combinations of SDSS and *GALEX* magnitudes yield similar trends with redshift. In general, the photometric errors on u are larger than for r, so we use the boundary of NUV $-$ r as a further cut on the large photometric catalogs. The lower limit of NUV $-$ r is 0 for the MIS spectroscopic sample and -0.3 for the AIS sample. This difference is likely a statistical fluctuation, and we have used -0.3 for all. The u-band errors in the confirmed QSO subset are all small, and so we may use this index as a redshift estimator in the photometric candidate catalogs provided we also eliminate objects with u-band errors of 0.5 mag or larger. Table 1 shows the data set sizes with these cuts.

3. PHOTOMETRIC SELECTION IN REDSHIFT BINS

From Figure 1, using cuts at $g-r=0.15$ or $r-i=0.1$, we can separate the QSOs into redshift bins roughly 0.3 to 0.7, 0.9 to 1.6, and >1.5. We restrict the overall ranges in these colors to -0.15 to 0.4, as almost all spectroscopically confirmed QSOs lie in this range.

The cleanest redshift cut is the bin 0.9 to 1.6, using $g-r$. There is contamination by a few low-redshift QSOs, and these can be eliminated by cutting the red end of the UV-optical index. This process includes 93% of the spectroscopically confirmed identified QSOs in this redshift range. Applying the same cuts to the entire MIS point source (MISP) sample, we find a total of 1646 QSO candidates, including 412 of spectroscopically known redshift. For the AIS point source (AISP) candidates, the cuts retain 82% of the spectroscopically confirmed known QSOs sample in the redshift range. Applying the same cuts to the entire AISP candidate sample, yields an extra 153 candidates, which scales to 187 allowing for this incompleteness. Since the total number is 623, the spectroscopic identification

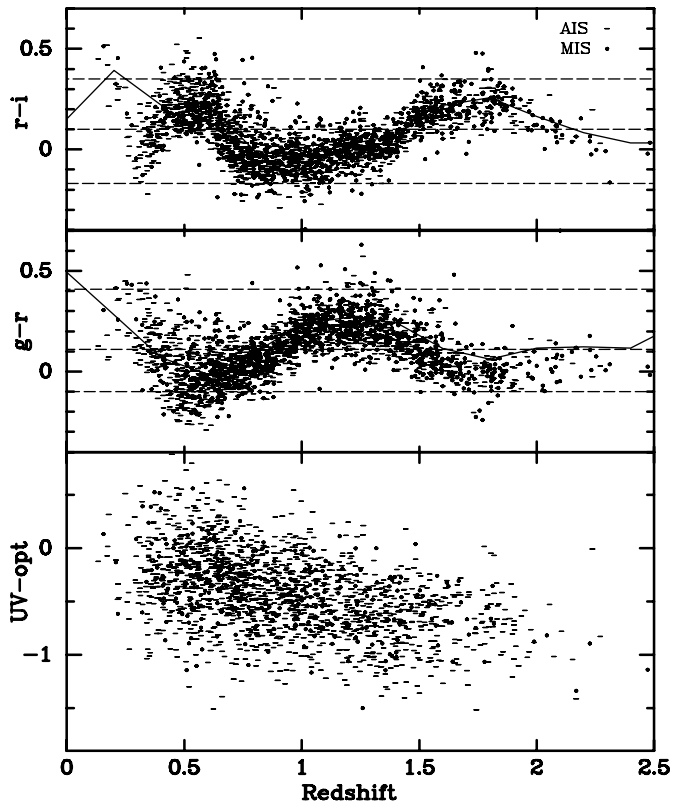

FIG. 1.—Colors with redshift for SDSS spectroscopic QSOs. The lines are models using standard QSO SEDs and the SDSS filters. The dashed lines indicate the color cuts used in our photometric samples.

of QSOs in the AISP sample appears to be complete at the 75% level.

The redshift range 0.9 to 1.6 is underrepresented in many samples because of the ground-based bandpasses and wavelengths of key emission lines. If we can separate the lower redshift QSOs (0.2 to 0.7), which have the same optical colors, we can derive new values for number counts and magnitudes for these redshift bins. The FUV − NUV color is very sensitive to this redshift range (see Fig. 7 of Bianchi et al. 2007). The $g - r$ and $r - i$ cuts on the known QSOs yield redshift bins above 1.5 and below 0.7, roughly. We can use the UV-visible color index to separate them, but there is enough scatter that we inevitably have some overlap. Thus, the high-redshift bin loses some candidates and accretes some low-redshift ones. The optimum cut for high-redshift QSOs yields close to the correct total numbers, but with some 10%–20% moved into and out of the redshift bin of interest. We looked at the mean magnitudes and magnitude distributions of the low- and high-redshift QSOs from the spectroscopically confirmed subsample, and find no significant differences (while the intermediate redshift objects do have a different distribution). Thus we consider that we can get lower limits to number counts but good magnitude distributions for low- and high-redshift candidates, although there will be some cross-contamination in the lists.

For the redshift greater than 1.5 objects, the MISP has 1120 objects in the candidates list, if the contamination average is the same. Of these, 200 have known (from SDSS spectra) redshifts. For the AISP sample, the numbers are 230 total candidates, of which 200 are spectroscopically confirmed. Thus, again we find almost complete identification of the QSOs in the AISP sample, in the SDSS spectroscopic database. This is mainly a reflection of the deeper limits in the MIS sample compared with the SDSS spectroscopic database. Figure 2 shows histograms of the magnitude distributions of these QSO candidates, compared with those for the spectroscopically confirmed QSO subsample, and the completeness limits of the spectroscopic data are evident. The fact that fewer faint candidates exist for the low- and high-redshift group (*top panels*) presumably is because there are few faint low-redshift objects, and the incompleteness of both the low- and the high-redshift sample due to our color selection discussed earlier.

Table 2 shows some mean magnitude and color indices for the different samples we discuss. The brighter limit of the AIS compared with the MIS is evident—also the difference between the spectroscopic sample and MIS limits. The color difference between the QSO candidates sample and spectroscopically confirmed QSO subsample are less obvious, and may indicate some differences in populations with redshift, as well as magnitude limits and errors, but may also be simply due to the different object selection criteria.

Among the sample of MIS and AIS extended source QSO candidates (which we call MISE and AISE) with spectra, we do not expect to find many QSOs, since they are normally registered as point sources. Indeed, there are only 34 spectroscopically

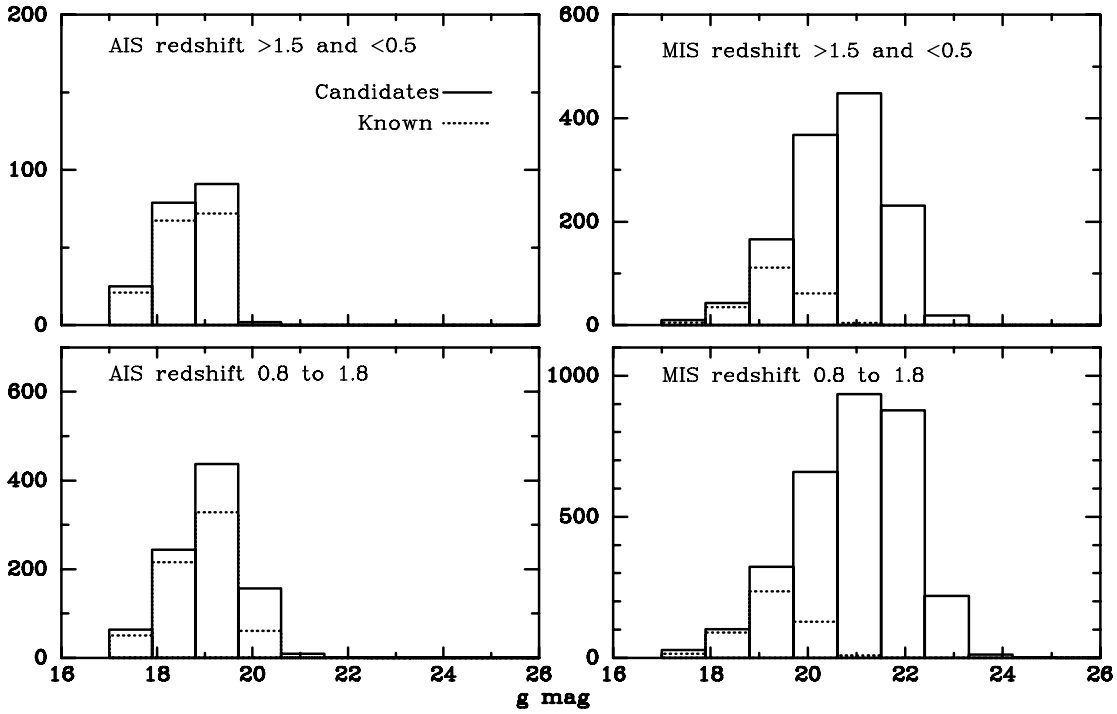

FIG. 2.—g magnitude distributions for the spectroscopically confirmed QSO samples and photometric QSO candidates, for the different catalogs as labeled.

TABLE 2

PROPERTIES OF POINT SOURCE SAMPLES

Property	MIS full all UV	MIS cand *gri* cuts	MIS QSOs spectra	AIS full all UV	AIS cand *gri* cuts	AIS QSOs spectra
Mean FUV mag	22.4	22.8	21.3	20.6	20.3	20.2
Mean g mag	20.9	20.8	19.2	19.2	19.0	18.8
Mean NUV $-$ r	0.76	0.95	0.87	0.44	0.80	0.62
Mean $g-r$	0.28	0.26	0.16	0.04	0.20	0.12
Mean z meas	...	...	1.14	...	...	0.95

confirmed QSOs in the MISE spectroscopic database, and 109 in the AISE. The same color cuts applied to these samples, reduce the QSOs counts to only 3, so this is clearly not going to produce any significant number of QSOs, at least to the magnitude limit of the SDSS spectra. This may not be true of the fainter sources, as we discuss in the next paragraph. The bright ones that are found have red colors and low redshift, indicating obscured nuclei, so that the host galaxies are more likely to be seen as resolved in the SDSS. While red QSOs may be a significant population (e.g., Hutchings et al 2006), this UV-optical database is not effective at finding them.

If we apply the $r-i$ and $g-r$ with NUV $-$ r color cuts to the MISE sample we find the total candidates shown at the bottom of Table 1. These are much larger numbers, so clearly they represent more than just QSOs. Further cuts to the $r-i$ and $g-r$ range correspond to redshifts 0.6 to 1.0: we get slightly fewer than the full range $g-r$ cut—some 4100 candidates. These numbers embody reasonable cuts to the g, i, and u-band errors, but are very sensitive to the error values for u-band. Overall, we feel these numbers suggest that of order 10% to 20% of the MISE sample may be faint QSOs. The mean g magnitudes of these samples are some 0.5 mag fainter than the MISP candidates (Table 2).

4. SPACE DENSITY AND COMPLETENESS

Figure 3 shows the number counts per square degree of sky for the QSO candidates we have discussed. It is notable that the counts do not match for the AIS and MIS at the bright end of the distributions. This indicates that there is incompleteness in the AIS list, presumably because of rejection by error bars in the weakest signal bands—u and z. The error bars for the MIS sample are much smaller because of their 15 times larger exposure times. In addition, the *GALEX* magnitude errors are larger for the AIS sample than the MIS, so there may be some systematic difference in the UV depth too. Overall, we thus consider that the MIS sample is complete to the maximum bin magnitudes ($g = 21$), and it falls off fast fainter than that. This diagram may be compared with Figure 11 of Bianchi et al (2007), for their broader selection of QSOs candidates.

Figure 4 shows the median magnitudes from each of the six bands, for the various QSO samples, together with the magnitude differences between them. We have compared a bright subset of the MIS samples with g magnitudes brighter than 19, for direct comparison with the AIS samples, which have very similar mean brightness. The two subsets are divided into redshift bins as Figure 3. In both sets, the bright MIS sample is redder than the AIS—i.e., it reaches fainter FUV magnitudes—while the r to z bands are brighter. This difference is more marked in the intermediate redshift set. The SED for the MIS total samples are essentially the same as the bright subsamples. This indicates that the longer exposures of the MIS gains more sensitivity with

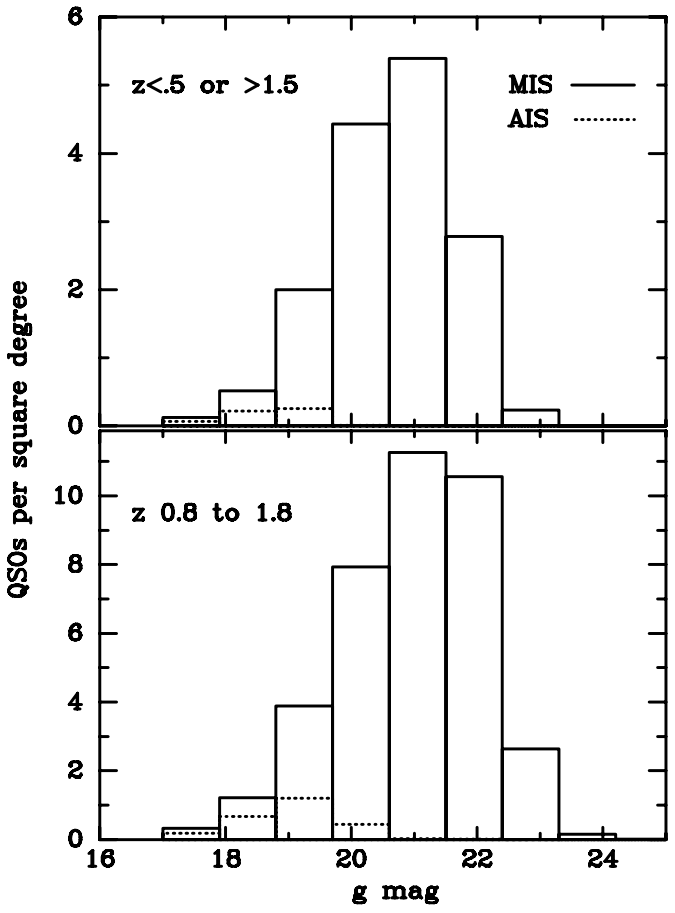

FIG. 3.—Space density of QSOs inferred from the photometric samples, as discussed in the text.

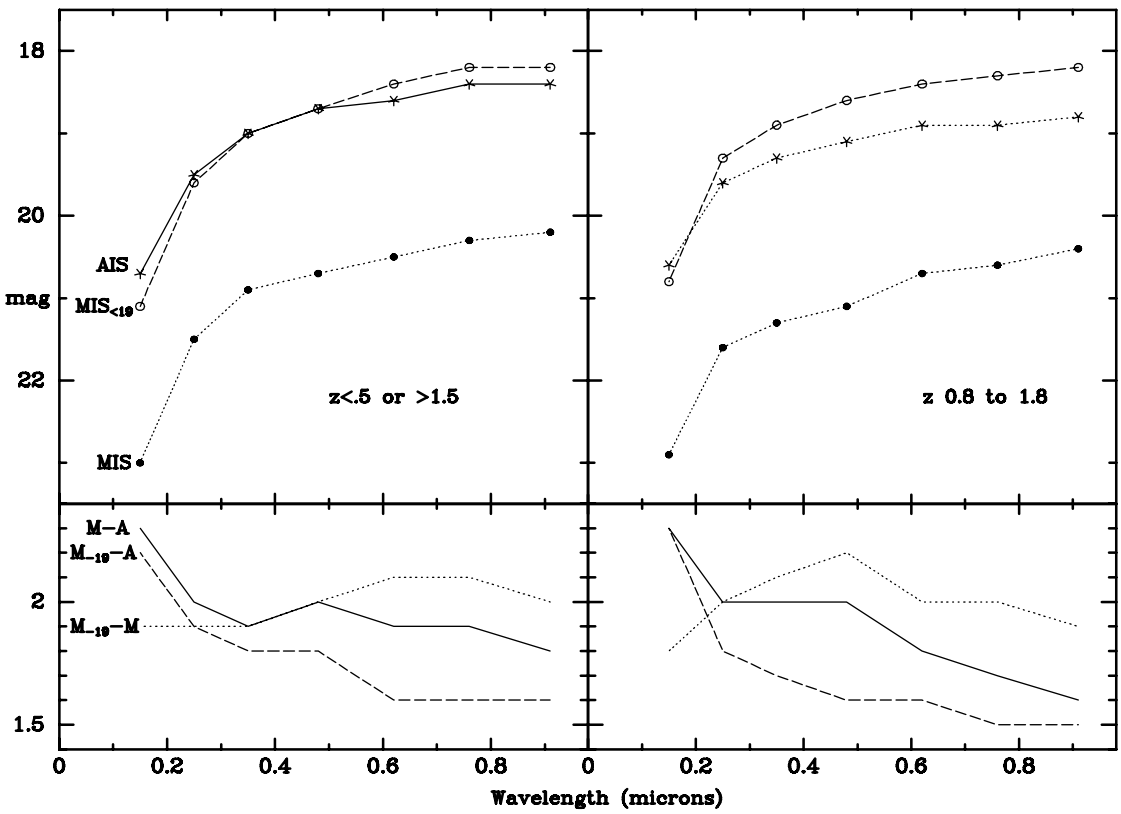

FIG. 4.—Magnitude SEDs and differences, from SDSS and *GALEX* filters for the photometric samples as labeled. The MIS QSOs are overall redder than the AIS.

respect to the AIS progressively as we go to shorter wavelengths. The color difference plots in the lower panels show the same things: the difference between MIS and AIS increases as we go to shorter wavelengths. This is consistent with the relative incompleteness of the AIS sample in Figure 3.

Note that the presence of strong emission at Lyα and C IV would make the median SED brighter in the NUV and u bands for the redshift 0.8 to 1.8 sample and the u band for the higher redshift sample. The Lyman edge drop would affect the FUV and NUV bands in the opposite sense. Allowing for these effects only increases the difference in MIS SED between the two redshift bins. The higher redshift QSOs are bluer, as may be expected for UV-bright rest frame SEDs.

5. DISCUSSION

The number counts of QSOs from the SDSS has been discussed by a number of authors (Richards et al 2004, 2005, 2006; van den Berk et al 2005; Schneider et al 2007) and their UV properties by Trammel et al 2007. The 2SLAC sample is discussed by Croom et al 2004. The UV selection of QSO candidates and their UV properties have been discussed by Bianchi et al 2007. We have compared our candidate counts with all these and find several points of interest, as they do not agree well. Figure 5 shows some comparisons. The MIS sample we have is different from the large sample of Richards et al, and their subsequent ones, in having lower space density peak, but more in the bright end of the distributions. Note, however, that the SDSS sample has a slightly lower density than the

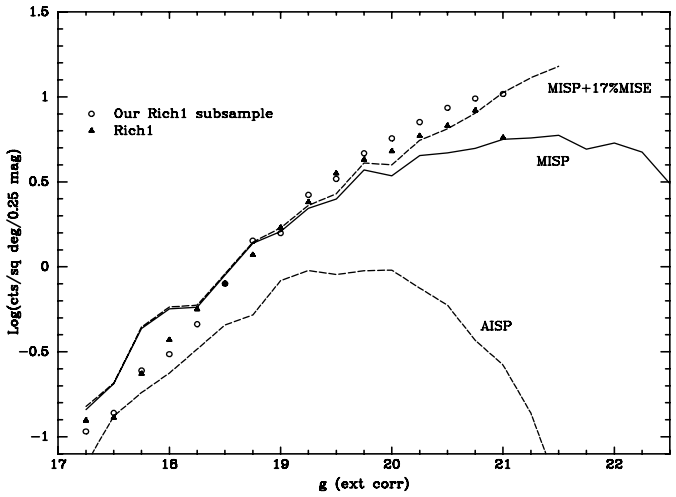

FIG. 5.—Space density distributions for our samples, compared with Richards et al 2004. The AIS sample is incomplete at all magnitudes. The difference between the full and our subsample of Richards points (see text) may show the errors inherent in samples of 70–80 deg^2.

GALEX-selected sample, when restricted to the same redshift range, even at bright magnitudes (Fig. 11 of Bianchi et al. 2007).

We have noted that our AIS sample gives lower source counts than the MIS. Looking at the magnitude 17–18 range, where our sample has an excess compared with the Richards et al candidates (and the Croom et al counts), a significant fraction of our QSO candidates have optical spectra, and for this fraction the confirmation rate is over 80% (Table 1; see also Bianchi et al. 2007 for more discussion). We believe that we can conservatively assume a figure of $>80\%$ for the purity of the entire candidate sample, as the contamination by stellar objects significantly decreases at fainter magnitudes. Therefore we consider our density estimates robust at this level, and the completeness fairly robust only for the redshift range around 1. It is possible that some of these are low luminosity QSOs or Seyferts, somehow excluded by the Richards et al selection. At the fainter end, if we assume that some of the MISE sample are misidentified QSO point sources, we boost mainly the faint end of the distribution, as discussed in the previous section.

One issue is that of the true sky coverage of the various samples discussed. It may be that they have been misestimated in some cases. Figure 5 shows our numbers from a small subset taken from the Richards et al catalog (about 65 deg^2, with some 3700 objects). It also shows the published values for the whole Richards sample, which suggests that their sky area is about right. Our subsample is similar to their total sample plots, but does have more faint sources. In all samples, various selections have been made in limiting magnitudes and magnitude errors, and even redshift in the case of the spectroscopic samples. A large unknown is the fraction of SDSS extended class sources that are misclassified point sources, as these potentially add many counts to the fainter end of the distributions. In our sample, the extended sources are by far the largest group. Richards et al claim that some 10% of sources fainter than $g\sim21$ are misclassified point sources. We find that 17% or more of them have colors consistent with QSOs. Figure 5 shows the distribution if we add this fraction of the MISE to the MISP. Overall, we suggest that the source counts are uncertain by a factor of the order of 2 at the faint end.

L. B. gratefully acknowledges NASA's support for construction, operation, and science analysis of the *GALEX* mission, developed in cooperation with the Centre National d'Etudes Spatiales of France and the Korean Ministry of Science and Technology.

REFERENCES

Bianchi, L., et al. 2007, ApJS, 173, 659
Bianchi, L. 2008, APSS, special issue, in press
Croom, S. M., Smith, R. J., Boyle, B. J., Shanks, T., Miller, L., Outram, P. J., & Loaring, N. S. 2004, MNRAS, 349, 1397
Hutchings, J. B., Cherniawsky, A., Cutri, R. M., & Nelson, B. O. 2006, AJ, 131, 680
Richards, G. T., et al. 2004, ApJS, 155, 257
Richards, G. T., et al. 2005, MNRAS, 360, 839
Richards, G. T., et al. 2006, AJ, 131, 2766
Schneider, D. P., et al. 2007, ApJ, in press, (arXiv:0704.0806)
Trammell, G. B., et al. 2007, AJ, 133, 1780
van den Berk, D. E., et al. 2005, AJ, 129, 2047