PHOTOMETRIC REDSHIFT PROBABILITY DISTRIBUTIONS FOR GALAXIES IN THE SDSS DR8

Erin S. Sheldon\textsuperscript{1}, Carlos E. Cunha\textsuperscript{2,3}, Rachel Mandelbaum\textsuperscript{4,5}, J. Brinkmann\textsuperscript{6}, and Benjamin A. Weaver\textsuperscript{7}

\textsuperscript{1}Brookhaven National Laboratory, Bldg 510, Upton, NY 11973, USA
\textsuperscript{2}Department of Physics, University of Michigan, 500 East University, Ann Arbor, MI 48109-1120, USA
\textsuperscript{3}Kavli Institute for Particle Astrophysics & Cosmology, Physics Department, and Stanford Linear Accelerator Center, Stanford University, Stanford, CA 94305, USA
\textsuperscript{4}Department of Physics, Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh, PA 15213, USA
\textsuperscript{5}Department of Physics, New York University, 4 Washington Place, New York, NY 10003, USA
\textsuperscript{6}Apache Point Observatory, P.O. Box 59, Sunspot, NM 88349, USA
\textsuperscript{7}Center for Cosmology and Particle Physics, Department of Physics, New York University, 4 Washington Place, New York, NY 10003, USA

Received 2011 December 12; accepted 2012 June 16; published 2012 August 2

ABSTRACT

We present redshift probability distributions for galaxies in the Sloan Digital Sky Survey (SDSS) Data Release 8 imaging data. We used the nearest-neighbor weighting algorithm to derive the ensemble redshift distribution \(N(z)\), and individual redshift probability distributions \(P(z)\) for galaxies with \(r < 21.8\) and \(u < 29.0\). As part of this technique, we calculated weights for a set of training galaxies with known redshifts such that their density distribution in five-dimensional color–magnitude space was proportional to that of the photometry-only sample, producing a nearly fair sample in that space. We estimated the ensemble \(N(z)\) of the photometric sample by constructing a weighted histogram of the training-set redshifts. We derived \(P(z)\)'s for individual objects by using training-set objects from the local color–magnitude space around each photometric object. Using the \(P(z)\) for each galaxy can reduce the statistical error in measurements that depend on the redshifts of individual galaxies. The spectroscopic training sample is substantially larger than that used for the DR7 release. The newly added PRIMUS catalog is now the most important training set used in this analysis by a wide margin. We expect the primary sources of error in the \(N(z)\) reconstruction to be sample variance and spectroscopic failures: The training sets are drawn from relatively small volumes of space, and some samples have large incompleteness. Using simulations we estimated the uncertainty in \(N(z)\) due to sample variance at a given redshift to be \(\sim 10\%–15\%\). The uncertainty on calculations incorporating \(N(z)\) or \(P(z)\) depends on how they are used; we discuss the case of weak lensing measurements. The \(P(z)\) catalog is publicly available from the SDSS Web site.

Key words: cosmology: observations – galaxies: distances and redshifts – large-scale structure of universe

Online-only material: color figures

1. INTRODUCTION

Photometric redshifts are estimates of redshift derived using broadband photometric observables such as magnitudes and colors (Baum 1962; Puschell et al. 1982; Koo 1985; Loh & Spillar 1986; Connolly et al. 1995). Typically, the set of observables for a given galaxy is not sufficient to uniquely specify its redshift, but only a probability distribution, the \(P(z)\). These \(P(z)\)'s are often relatively broad. For simplicity of use and interpretation, one commonly uses a single number, the photometric redshift, as the best estimate of a galaxy’s redshift. As several recent works have shown (Mandelbaum et al. 2008; Cunha et al. 2009; Wittman 2009; Bordoloi et al. 2010; Abrahamse et al. 2011), the use of a single number to represent the photometric redshift, as the best estimate of a galaxy’s redshift, is necessary to uniquely specify its redshift. Working with the full \(P(z)\) for each galaxy yields better estimates of the overall redshift distribution, \(N(z)\), and can decrease biases in cosmological analyses. We note that several photo-z codes exist that can produce a \(P(z)\) per galaxy, e.g., \textit{Le Phare} (Arnouts et al. 1999; Ilbert et al. 2006), \textit{ZEBRA} (Feldmann et al. 2006), \textit{BPZ} (Coe et al. 2006), \textit{ArborZ} (Gerdes et al. 2010), and our own method (Cunha et al. 2009), henceforth referred to as ProbWTS, which is an acronym for probability distributions from weighted training sets. We use \(P(z)\) when referring to the \(P(z)\) derived from ProbWTS.

In this paper, we describe a \(P(z)\) catalog for objects detected in the Data Release 8 (SDSS DR8; Aihara et al. 2011) of the Sloan Digital Sky Survey III (SDSS III; Eisenstein et al. 2011). We use the method of Cunha et al. (2009), which was also applied to SDSS DR7 (Abazajian et al. 2009), with improvements in the training set and photometry. The DR7 catalog of Cunha et al. (2009) has been successfully used in cosmological analyses, allowing, for example, for the first measurement of the transverse baryon acoustic oscillation (BAO) scale derived purely from angular information, i.e., without using the three-dimensional power spectrum (Carnero et al. 2012) and for the measurement of the growth of structure using photometric luminous red galaxies (LRGs; Crocce et al. 2011).

This paper is organized as follows. In Section 2, we discuss the method and in Sections 3–5 we describe the data and sample selection. In Section 6, we discuss the training set and in Sections 7 and 8 we show our results, including information about their public release, and estimate errors. In Section 9, we discuss the proper usage of these results. As an example, in Section 10 we discuss the particular case of weak gravitational lensing calculations, and explore the expected errors on such a measurement. Finally, in Section 11 we summarize our results.

2. METHOD

The algorithm is detailed in Lima et al. (2008) and Cunha et al. (2009). The method is to derive weights for a training set of spectroscopically confirmed galaxies such that the distribution of relevant quantities, such as magnitudes or colors, matches that of a set of galaxies without known redshifts, henceforth the photometric sample. Assuring these quantities correlate
with redshift, and are the only relevant quantities for redshift
determination, the resulting weighted redshift histogram is
proportional to the redshift probability distribution \( N(z) \) of the
photometric sample.

The weighting forces the distributions of observables of
the two samples to be proportional, essentially creating a
“fair sample” from the training set, avoiding errors due to
population differences. An additional advantage is that the
technique naturally identifies the regions of intersection in the
observable space between training and photometric samples.
Only in the intersection one can safely use the training sample to
make inferences about the photometric sample.

The method, as well as any other training-set-based estimator,
will not work if there are redshift selection issues localized in
the space of observables. For example, consider a sample of
galaxies occupying the same region of observable space. If there
is a systematic selection such that the subset of the galaxies with
spectra has a systematically different redshift distribution from
the rest of the galaxies, the weights will not be able to recover
the redshift distribution correctly. Note that if redshifts cannot
be obtained for a galaxies with a particular type of spectral
energy distribution (SED), there will only be a bias if the redshift
distribution of galaxies with that SED at that particular region
of observable space does not match that of the other galaxies.
In general, this does not appear to be a dominant issue for this
data set, but may affect some particular populations.

2.1. Nearest-neighbor \( P(z) \) Redshift Estimators

2.1.1. Weights

In this section, we briefly review the weighting method\(^8\)
of Lima et al. (2008), which is required for computing \( P(z) \).
We define the weight, \( w_\beta \), of a galaxy in the spectroscopic
training set as the normalized ratio of the density of galaxies
in the photometric sample to the density of training-set galaxies
around the given galaxy. These densities are calculated in a local
neighborhood in the space of photometric observables, e.g.,
multi-band magnitudes. In this case, the SDSS ugriz magnitudes
are our observables; in practice we use four colors and the \( r \)-band
magnitude. The hypervolume used to estimate the density is set
to be the Euclidean distance of the galaxy to its 100th nearest-
neighbor in the training set. Note, when the training data are
assembled from multiple spectroscopic samples, we estimate the
weights from the entire combined sample rather than separately
from individual samples.

The weights can be used to estimate the redshift distribution
\( N(z) \) of the photometric sample:

\[
N(z) = \sum_{\beta=1}^{N_B} w_\beta \delta(z - z_\beta).
\]

(1)

For a bin \( z_1 < z < z_2 \), we sum the weights, \( w_\beta \), of all training-
set galaxies that have redshift \( z_\beta \) fall within that bin. Lima
et al. (2008) and Cunha et al. (2009) show that this indeed
provides a nearly unbiased estimate of the redshift distribution
of the photometric sample, \( N(z) \), provided the differences
in the selection of the training and photometric samples are
solely in the observable quantities used to calculate the weights.
For example, if the photometric sample has a morphology-
dependent cut, the same cut should be applied to the training
sample or morphology should be one of the observables used to
measure weights.

2.1.2. \( P(z) \)

To estimate the redshift error distribution for each galaxy,
\( P(z) \), we adopt the method of Cunha et al. (2009). The \( P(z) \)
for a given object in the photometric sample is simply the redshift
distribution of the \( N \) nearest neighbors in the training set:

\[
\hat{P}(z) = \frac{N_T}{\sum_{\beta=1}^{N_T} w_\beta \delta(z - z_\beta)}.
\]

(2)

This expression is the same as Equation (1) but is limited to
the nearest neighbors of a given object. We choose \( N_T = 100 \)
for this study, and estimate \( P(z) \) in 35 redshift bins between
\( z = 0 \) and 1.1. We can also construct a new estimator for \( N(z) \)
by summing the \( \hat{P}(z) \) distributions for all \( N \) galaxies in the
photometric sample,

\[
N(z) = \sum_{i=1}^{N_{Gal}} \hat{P}_i(z).
\]

(3)

A key difference between the estimators in Equations (1) and (3)
is that the hyperball used to select the nearest neighbors is
centered on a training-set object in the weights estimator, but
centered on a photometric set object for the \( P(z) \) estimator.
The estimators of Equations (1) and (3) agree in the limit of
very large training sets, but Equation (3) is subject to biases
otherwise. For training sets smaller than tens of thousands
of galaxies, one can improve the \( P(z) \) by multiplying each \( P(z) \)
by the ratio of \( N(z) \) to \( N(z) \). That is,

\[
P(z) \rightarrow P(z) \frac{N(z)}{N(z)}.
\]

(4)

This correction essentially corresponds to using the weights
estimate as a prior on the \( P(z) \).

3. PHOTOMETRIC DATA

The photometric data were drawn from DR8 of the SDSS III.
Full details are given in the data release paper (Alhara et al.
2011). As compared with the earlier DR7 release (Abazajian
et al. 2009), DR8 includes an additional 2500 deg\(^2\) of new
imaging in the southern galactic cap (SGC), acquired to facil-
itate spectroscopic target selection for the Baryon Oscillation
Spectroscopic Survey (BOSS), which is part of SDSS III.
SDSS (York et al. 2000) images are gathered using the 2.5 m
at Apache Point (Gunn et al. 2006) with the camera (Gunn et al.
1998) running in time-delay-and-integrate mode. Observations
are taken in each of the SDSS bandpasses (ugriz; Fukugita
et al. 1996) nearly simultaneously as sky moves across bands
in the order riecz. The data were taken during photometric
nights under relatively good seeing conditions (Hogg et al.
2001). A series of pipelines are run to calibrate the data
(Padmanabhan et al. 2008; Smith et al. 2002; Tucker et al.
2006), derive astrometry (Pier et al. 2003), and calculate fluxes,
shapes, and other interesting quantities (Lupton et al. 2001).
Note the calibrations used for these data are derived using the
“ubercalibration” technique presented in Padmanabhan et al.
(2008).

\(^8\) The weights and \( P(z) \) codes are available at
http://slac.stanford.edu/~ccunha/nearest/ Alternatively, the code can be
accessed at the git repository https://github.com/dcurtis147/probets
4. PHOTOMETRIC QUANTITIES

In this section, we describe the photometric quantities used in the creation of the input catalog. Most of these quantities are measured by the SDSS photometric pipeline PHOT. An early version of the pipeline is described in Lupton et al. (2001); other details can be found in the SDSS data release papers, e.g., Adelman-McCarthy et al. (2006) and at the SDSS III Web site.9 We give a few additional details below. In comparison to DR7, the DR8 makes use of an updated version of the PHOT software reduction pipeline, v5.6 rather than v5.4, including some updates to sky subtraction that can change galaxy photometry and, potentially, the $P(z)$’s.

For colors we use the SDSS “model magnitudes,” which we refer to as $c_{\text{model}}$. Each object is fit to an elliptical exponential disk and an elliptical De Vaucouleurs’ profile convolved with a double Gaussian approximation to the point-spread function (PSF) model interpolated to the location of the object (Lupton et al. 2001; Sheldon et al. 2004). For the $c_{\text{model}}$, the best-fit model in the $r$ band is then used to extract the flux in the other four bandpasses, accounting appropriately for the PSF in each band. Thus, the effective aperture is the same for all bands, which is appropriate for extraction of color information.

We use “composite model magnitudes” as an approximate total magnitude for each object, which we refer to as $c_{\text{model}}$. For each bandpass separately, PHOT does an additional joint fit to a non-negative linear combination of the best-fitting exponential and De Vaucouleurs’ models. This fit determines an additional parameter $\text{FRAC}_{\text{DEV}} (f_{\text{dev}})$, which is the fraction of the total flux estimated to come from a De Vaucouleurs’ profile. The composite model flux in each band is then

$$\text{Flux}_{c_{\text{model}}} \equiv (1 - f_{\text{dev}}) \times \text{Flux}_{\text{exp}} + f_{\text{dev}} \times \text{Flux}_{\text{dev}}.$$  \hspace{1cm} (5)

Because this procedure is carried out separately per band, the effective aperture for each band is different, so these magnitudes are not appropriate for estimating colors.

For quality assurance, we use bits from the OBJECT bitmask output by PHOT.11 We also use the RESOLVE_STATUS flag to choose primary observations.12 We will describe how the flags are used in Section 5.

5. PHOTOMETRIC SAMPLE SELECTION

5.1. Star–Galaxy Separation

The PHOT pipeline uses the concentration $c$ to separate stars from galaxies. The concentration is the difference between magnitude determined from the best-fitting PSF model $p_{\text{sfm}}$ and the $c_{\text{model}}$ which is the better fitting of the exponential and De Vaucouleurs’ models convolved with the local PSF:

$$c \equiv \text{psfm} - c_{\text{model}}.$$  \hspace{1cm} (6)

For stellar objects, the scale of the $c_{\text{model}}$ approaches a delta function and the result becomes equivalent to the $p_{\text{sfm}}$. Thus the concentration should be $\geq 0$ within the noise, with stars close to zero and galaxies greater than zero. The pipeline defines galaxies as objects with $c > 0.145$ where $c$ is derived from the summed fluxes from all bandpasses.13

At our magnitude limits, the stellar contamination is relatively large. Using a small, space-based, high angular resolution data set matched to SDSS data as a truth table, the approximate stellar contamination can be determined. At $r = 21$, the contamination is a few percent, but the contamination increases to approximately 10% at $r = 22$.14

For studies where completeness and purity must be known precisely, Scranton et al. (2005) recommend using probabilistic star–galaxy separation at fainter mags ($r > 21$); i.e., attempt to determine the probability that an object is a galaxy and either use that as a weight or make appropriate cuts.

In practice, the end user should choose a subset of the data that suits their needs. We provide a catalog here that should be a superset of objects that can be further trimmed.

5.2. Other Cuts

We remove objects for which the extinction-corrected (Schlegel et al. 1998) model flux is not well determined, in at least one of the photometric bands, by demanding $g \leq 21 \quad (g \leq 22 \quad | r < 22 | \quad i < 20.5 \quad | | z < 20.1)$, where $| |$ is a C-style “or” meaning any one of the criteria should be true. This cut removes spurious entries in the catalog. As such, this cut is relatively unimportant compared to the cut in $r$ given below.

We additionally demand a detection in both the $r$ and $i$ bands. Rather than applying a magnitude cut, we instead use the OBJECT processing flags BINDED{1,2,4}, which indicate the object was detected in the original image (binned by 1), the $\times 2$ binned image, or the $\times 4$ binned image, respectively (Stoughton et al. 2002).

We remove all objects that have the following OBJECT flags set: SATUR, BRIGHT, DEBLEND TOO MANY PEAKS, PEAKCENTER, NOTCHECKED, and NOPROFILE as well as objects that are (BLENDED & NODBLEND); in other words, detected to be blended but not successfully deblended into components (where && is a C-style “and,” meaning both criteria are required to be true).

We demand that the data are photometric in each band, as indicated by the CALIB_STATUS flag PHOTOMETRIC.

We only use objects marked as SURVEY_PRIMARY in their RESOLVE_STATUS flags field. Different scans on the sky image the same objects due to the small overlap regions between adjacent scans, overlaps at the end of the scan lines where the great circles converge, and re-observed scan lines. This results in duplicate observations for many objects. These duplicates are “resolved” and only a single observation is assigned SURVEY_PRIMARY. Note that the SURVEY_PRIMARY flag also implies that, if the object is blended, it is either a child or not deblended further. This cut is made in the OBJECT flags as ![BRIGHT && ![BLENDED || NODEBLEND || nchild == 0).  

We also require the extinction-corrected (Schlegel et al. 1998) $c_{\text{model}}$ in the $r$ band to be in the range [15.0, 21.8]. This cut is more stringent than our initial magnitude cut, which just demanded a good detection in at least one of the bands and therefore might allow galaxies fainter than 22nd magnitude in $r$ into the sample. Also this cut is in $c_{\text{model}}$ rather than $c_{\text{model}}$. We also restrict the extinction-corrected $c_{\text{model}}$ to be within the range [15.0, 29.0] for all bands in order to ensure reasonable colors for the galaxies.

We make broad geometrical cuts on the catalog. We trim the objects to the BOSS footprint, shown in Figure 1. We also remove any objects near stars in the tycho2 catalog (Høg et al.

---

9 http://www.sdss3.org
10 http://www.sdss3.org/dr8/algorithms/magnitudes.php
11 http://www.sdss3.org/dr8/algorithms/flags_detail.php
12 http://www.sdss3.org/dr8/algorithms/resolve.php
13 http://www.sdss3.org/dr8/algorithms/classify.php
14 http://www.sdss.org/DR7/products/general/stargalsep.html
using a variable radius that depends on the magnitude of the star:

\[ \text{radius} = (0.0802 \times B_T^2 - 1.860 \times B_T + 11.625)/60.0, \]

where \( B_T \) is the Tycho magnitude and \( r \) is in degrees (Blanton et al. 2005). Finally, we remove all objects from images taken where a \( u \) amplifier was not working.\(^{15}\)

The final photometric catalog contains 58,533,603 objects. The distributions of extinction-corrected \( r \)-band \( cmodel\text{mag} \) and colors derived from extinction-corrected \( model\text{mag} \) are shown in Figure 2.

The following list shows a breakdown of the fraction of objects lost to a given cut after applying \( r < 21.8 \) and geometrical cuts, and when only that cut is applied:

1. \text{BINED flag} cuts: 0.92.
2. \text{CALIB\_STATUS flag} cuts: 0.99.
3. \text{OBJECT flag} cuts: 0.98.
4. Cutting all \( model\text{mags} \) to \([15.0, 29.0]\): 0.80. Independently by band: \( u \): 0.82; \( g \): 0.99; \( r \): 1.00; \( i \): 0.99; \( z \): 0.98.

The loss of 20\% of the galaxies to the last cut is primarily to the \( u \) limit. This means the sample is not strictly flux limited to \( r < 21.8 \). Star-forming galaxies will be retained more so than those with older stellar populations.

6. TRAINING SAMPLES

We use a spectroscopic training set drawn from a number of sources. These sources contain mostly galaxies and a small number of stars in order to help characterize stellar contaminants from the photometric sample at low redshift. In the following sections, we give short details on each sample and describe our process for matching to the photometric sample.

6.1. Samples Used in this Study

1. 435,878 redshifts from the SDSS spectroscopic samples, principally from the MAIN (Strauss et al. 2002) and LRG (Eisenstein et al. 2001) samples, with confidence level \( z\text{conf} > 0.9 \) and \( r \)-band \( cmodel\text{mag} < 19.5 \).

2. 445 objects from the Canadian Network for Observational Cosmology (CNOc) Field Galaxy Survey (CNOc2; Yee et al. 2000)\(^{16}\) with \( R\text{val} > 4 \) for \( Sc = 2 \) or 4, or \( R\text{val} > 5 \) for \( Sc = 5 \).

3. 151 from the Canada–France Redshift Survey (CFRS; Lilly et al. 1995)\(^{17}\) with \( Q\text{lass} \geq 3 \).

4. 1868 from the Deep Extragalactic Evolutionary Probe 2 survey (DEEP2; Weiner et al. 2005)\(^{18}\) with \( z\text{qual} \geq 3 \). Of these, 1499 are an approximately magnitude-limited sample from the Extended Groth Strip (EGS). The remainder is \( BRI \) color selected to target \( z > 0.7 \) galaxies, hereafter denoted the non-EGS sample.

5. 197 from the Team Keck Redshift Survey (TKRS; Wirth et al. 2004)\(^{19}\) with \( Q = 4 \) and \( Q = -1 \). The flag \( Q = -1 \) corresponds to stars, and only two of them were left in the final sample.

6. 8633 LRGs from the 2dF-SDSS LRG and QSO Survey (2SLAQ; Cannon et al. 2006)\(^{20}\) with \( qop \geq 3 \).

7. 2080 from \( z\text{COSMOS} \) redshift survey Lilly et al. (2007), with \( cc = 3.4 \) \( \leq 3.5 \) \( \leq 4.4 \) \( \leq 4.5 \) \( \leq 9.5 \). Note that 2046 galaxies had \( cc = 3.5 \) \( \leq 4.5 \), and 24 had \( cc = 9.5 \).

8. 1587 from the VIMOS VLT-Deep Survey (VVDS; Garilli et al. 2008)\(^{21}\) with \( z\text{qual} = 3 \) \( \leq 4 \).

9. 16,874 from four fields of the PRIMUS survey (PRIMUS; Coil et al. 2011; R. Cool et al. 2012, in preparation).\(^{22}\) Only PRIMUS objects with \( Q = 4 \) were used.

In Table 1, we present some statistics about each training set.

6.2. Matching to SDSS Imaging Data

We spatially match the training sets listed in Section 6.1 to the photometric catalog described in Section 5. We choose the closest match within 2\'. By performing this match, we place

\(^{15}\) http://www.sdss.org/dr7.1/start/aboutdr7.1.html\#imcaveat

\(^{16}\) http://www.astro.toronto.edu/~cnoc/cnoc2.html

\(^{17}\) http://www.oamp.fr/people/tresse/cfسم/cfrs.html

\(^{18}\) http://deep.berkeley.edu/DR3

\(^{19}\) http://tksrvr.keck.hawaii.edu/tksurvey/

\(^{20}\) http://www.2slaq.info/

\(^{21}\) http://www.oamp.fr/virmos/vvds.htm

\(^{22}\) http://cass.ucsd.edu/~acoil/primus/
Figure 2. Distributions of photometric quantities for the photometric sample and training sample. The upper left panel shows the extinction-corrected $r$-band \textit{cmodelmag}. Both samples are cut at $r < 21.8$. Also shown is the weighted histogram for the training sample where the weights are derived to produce distributions approximately proportional to the photometric sample. The following four panels show extinction-corrected colors based on \textit{modelmag}. The bottom right panel shows the distribution of the derived weights for the training sample. Note the weight calculation is performed in the full five-dimensional space; we show the projections here to help visualization.

(A color version of this figure is available in the online journal.)

Table 1  

| Survey       | Number of Objects | Area (sq. deg.) | Weight Fraction |
|--------------|-------------------|-----------------|-----------------|
| PRIMUS$^*$   | 16874             | 5.2             | 0.63            |
| zCOSMOS$^*$  | 2080              | 1.7             | 0.075           |
| SDSS DR5     | 435875            | 5740            | 0.074           |
| 2SLAQ        | 8633              | 180             | 0.060           |
| VVDS$^*$     | 1587              | 4.0             | 0.060           |
| DEEP2-EGS$^*$| 1499              | 0.4             | 0.058           |
| SDSS DR5 ($r < 17.8$) | 376625       | 5740            | 0.017           |
| CNOC2$^*$    | 445               | 0.4             | 0.016           |
| DEEP2-nonEGS$^*$ | 369           | 2.8             | 0.014           |
| CFRS$^*$     | 151               | <0.1            | 0.0076          |
| TKRS$^*$     | 197               | 0.07            | 0.0055          |

Notes. Number of galaxies, area in square degrees, and fractional contribution to the weights estimate of $N(z)$. The "$^*$" indicates samples that are approximately flux limited to our selection depth.

7. RESULTS

We use the algorithm described in Section 2 to derive weights for each training-set galaxy. We then use these weights to calculate a weighted redshift histogram, which, under our assumptions, should be proportional to that of the photometric set. We also derive individual redshift probability distributions $P(z)$ for each photometric galaxy.
The Astrophysical Journal Supplement Series, 201:32 (12pp), 2012 August

Sheldon et al.

7.1. Derived Weights in Observable Space

The $r$-band $c_{\text{modelmag}}$ and colors based on $\text{modelmag}$ for the photometric and training sets are shown in Figure 2. Also shown are the derived weights for the training set and the resulting weighted histograms. These are the fundamentally new calculations presented in this work.

The weighted training-set distributions should be approximately proportional to the photometric set distributions in order to derive good redshift distributions. There are deviations at $g - r \sim 1.5$ and $r - i \sim 0.6$, but qualitatively the distributions are close. We focus on the accuracy of the recovered redshift distributions rather than a detailed comparison of these distributions.

7.2. Derived $N(z)$

In Figure 3, we present the recovered redshift distribution for the entire sample as described in Section 5. Also shown is the redshift distribution of the original training set. These distributions are in qualitative agreement with those shown in Cunha et al. (2009), although that sample had a fainter $r$-mag limit at 22.0. Note the sub-plot showing the region near $z = 0$. As expected, there is a non-zero fraction of the overall distribution near redshift zero. The fraction of the probability at $z < 0.002$ is about 0.4%. It is not known exactly how many stars are in the photometric sample, but this is probably a lower limit on the stellar contamination (see Section 5.1). We will estimate the errors on this distribution in Section 8. These $N(z)$ data are presented in Table 2.

![Figure 3. Reconstructed redshift distribution for SDSS galaxies with $r < 21.8$. The overall reconstructed distribution, shown in red, is derived by creating a weighted histogram of the training-set redshifts as described in the text. Also shown in magenta is the sum of all $P(z)$ derived for individual galaxies. The unweighted training-set redshift distribution is shown in blue. The expected errors on these distributions from cosmic variance in the training set are shown in Figure 7. The excess at $z \sim 0$ is due to stars in training set having significant weight; more detail at low redshift is shown in the inset. This excess is at least partly due to the presence of real stars in our photometric sample resulting from imperfect star-galaxy separation. The fraction of the distribution at $z < 0.002$ is 0.4%, which is probably a lower bound on the stellar contamination. (A color version of this figure is available in the online journal.)](image)

| $z_{\text{min}}$ | $z_{\text{max}}$ | $N(z)$ | Sample Variance Error |
|------------------|------------------|--------|-----------------------|
| 0.000            | 0.031            | 0.150  | 0.052                 |
| 0.031            | 0.063            | 0.822  | 0.215                 |
| 0.063            | 0.094            | 1.837  | 0.409                 |
| 0.094            | 0.126            | 2.815  | 0.503                 |
| 0.126            | 0.157            | 3.909  | 0.509                 |
| 0.157            | 0.189            | 5.116  | 0.725                 |
| 0.189            | 0.220            | 6.065  | 0.905                 |
| 0.220            | 0.251            | 6.477  | 0.767                 |
| 0.251            | 0.283            | 6.834  | 0.817                 |
| 0.283            | 0.314            | 7.304  | 0.868                 |
| 0.314            | 0.346            | 7.068  | 0.645                 |
| 0.346            | 0.377            | 6.771  | 0.785                 |
| 0.377            | 0.409            | 6.587  | 0.609                 |
| 0.409            | 0.440            | 6.089  | 0.627                 |
| 0.440            | 0.471            | 5.165  | 0.602                 |
| 0.471            | 0.503            | 4.792  | 0.522                 |
| 0.503            | 0.534            | 4.228  | 0.383                 |
| 0.534            | 0.566            | 3.664  | 0.394                 |
| 0.566            | 0.597            | 3.078  | 0.364                 |
| 0.597            | 0.629            | 2.604  | 0.275                 |
| 0.629            | 0.660            | 2.130  | 0.224                 |
| 0.660            | 0.691            | 1.683  | 0.191                 |
| 0.691            | 0.723            | 1.348  | 0.156                 |
| 0.723            | 0.754            | 0.977  | 0.141                 |
| 0.754            | 0.786            | 0.703  | 0.102                 |
| 0.786            | 0.817            | 0.521  | 0.080                 |
| 0.817            | 0.849            | 0.339  | 0.060                 |
| 0.849            | 0.880            | 0.283  | 0.048                 |
| 0.880            | 0.911            | 0.187  | 0.037                 |
| 0.911            | 0.943            | 0.141  | 0.031                 |
| 0.943            | 0.974            | 0.104  | 0.027                 |
| 0.974            | 1.006            | 0.081  | 0.020                 |
| 1.006            | 1.037            | 0.055  | 0.017                 |
| 1.037            | 1.069            | 0.043  | 0.015                 |
| 1.069            | 1.100            | 0.034  | 0.012                 |

Notes. Reconstructed redshift distribution $N(z)$ for SDSS galaxies with $r < 21.8$. The first two columns specify the redshift range of the bin and the third is the reconstructed $N(z)$, with arbitrary normalization. The fourth is the sample variance errors on $N(z)$ derived from simulations, which we expect to be the dominant uncertainty. These sample variance errors should be thought of as a rough estimate. A more perfect match would require a simulation more specifically tuned to the SDSS data.

7.3. Derived $P(z)$

Also shown in Figure 3 is the summed $P(z)$ derived for individual galaxies. The uncorrected $N(z)_{\text{uni}}$ is, characteristically, slightly more peaked than $N(z)_{\text{wei}}$. In Section 7.3.1, we apply Equation (4) to correct the $P(z)$s.

In Figure 4, we show six randomly chosen $P(z)$s. Each panel contains a $P(z)$ drawn from a particular magnitude range in extinction-corrected $r$-band $c_{\text{modelmag}}$; these ranges are given in the figure caption. This figure captures the general trend that the $P(z)$s are broader at fainter magnitudes, which is the expected behavior.

The uncertainty in individual $P(z)$s is typically dominated by shot-noise error. The scale of both statistical and systematic uncertainties in the individual $P(z)$s is strongly correlated with the width of the $P(z)$ (Cunha et al. 2009). A broader $P(z)$ reflects a larger degeneracy in observable space, and requires more training-set objects to characterize. Figure 5 shows the
distribution of objects in the photometric sample as a function of $r$-band magnitude and $1\sigma$ width of the $P(z)$. We define $1\sigma$ as

$$\sigma^2 = \frac{1}{N} \sum_{i} (\langle z \rangle - z_i)^2,$$

where $\langle z \rangle$ is the mean redshift and the sum is over the $N$ training-set galaxies used to construct the $P(z)$. The contours indicate factor of two changes in density.

We recommend using the $1\sigma$ or other width measures of the $P(z)$ as the most efficient way to trim the sample for improved precision and accuracy. The $P(z)$ width should also be a reasonable error estimator for use with other photo-$z$ methods. However, we discourage using the peak or some other single number statistic derived from the $P(z)$ as a proxy for redshift. See Section 9 for more details.

### 7.3.1. Correction to $P(z)$

As we will demonstrate in Section 10, the individual $P(z)$s are somewhat less accurate than the overall $N(z)$. We can correct the individual $P(z)$ to agree, in the mean, with the overall $N(z)$ using Equation (4). This correction factor is shown in Figure 6. At $z \gtrsim 0.9$ neither the $N(z)$ or the
summed $P(z)$ is well constrained, and the correction factor is noisy. For $z > 0.9$, we use the average correction from that range.

### 7.4. Differences from Previous $P(z)$ Derived Using this Method

Unlike for the DR7 catalog, we did not use repeat observations of our training-set galaxies. The use of repeats can provide more localized and smoother $P(z)$ estimates, and are often useful. However, because only part of our sample had repeat observations, the use of repeats would effectively increase the sample variance of our results. The use of repeats may be beneficial for LRGs because the training set is not sample variance limited in this case. We may release a catalog trained on repeat observations at a future date.

### 7.5. Acquiring the Data

The $P(z)$s for all galaxies are available from the SDSS III Web site. The data are available in both FITS format and ASCII. The objects are split into different files according to their SDSS run id, with each row in the file representing the data for a single SDSS object. The data for each object are SDSS id, the input colors and magnitude for each object, equatorial latitude and longitude, and the estimated $P(z)$.

### 8. SOURCES OF ERROR

Extensive tests of the ProbWTS method have been performed on real data by Lima et al. (2008) using a sample very similar ours. Further tests using simulations were performed by Cunha et al. (2009). In this section, we discuss in detail a number of possible sources of error.

As detailed in Cunha et al. (2009), the derived weights, and inferred $N(z)$, are susceptible to at least four kinds of training-set selection effects: spectroscopic failures, two types of large-scale structure bias (sample variance + shot noise in the training set), and selection in non-photometric observables. In addition, the fact that the weights use a non-infinitesimal volume in color–magnitude space to re-weight the photometric set can yield a small Eddington bias to the recovered distribution. And, as mentioned previously, incorrect star–galaxy separation can result in incompleteness and contamination of the sample. Because our training set consists of many different surveys with different characteristics, it is important to quantify the contribution of each to the overall result. Table 1 lists, for each of the surveys comprising the training set, the number of objects, the approximate area, and the fraction the survey contributes to the weighted estimate of the overall redshift distribution. This fraction is calculated by summing the weights assigned to objects in each survey and dividing by the sum of weights from the entire training set.

From Table 1, we see that PRIMUS carries the most weight by a large margin at 62%. Overall, the magnitude-limited surveys that reach our selection depth of $r = 21.8$—PRIMUS, TRKR, CNOC2, DEEP2-EGS, CFRS, VVDS, and zCOSMOS—represent about 81% of the total weight. This is desirable, because it minimizes the risk of bias in our assessment of errors in what follows. The table also shows that the SDSS MAIN sample ($r < 17.8$) contributes only 1.7% of the weights, which is consistent with the fraction expected from simulations for a flux-limited sample to $r < 21.8$. The remainder of the SDSS

#### Figure 7

Top panel: simulated redshift distribution with errors for an $r < 21.8$ sample. The error bars are the 1σ simulated variability due to sample variance in the catalogs comprising the training set. Also shown is the estimated $N(z)$ for our sample. Lower panel: estimated $N(z)$ combined with the predicted sample variance errors from the simulation.

(A color version of this figure is available in the online journal.)

spectra are LRGs to $r < 19.4$, which make a contribution to the total weight at 7.4%.

In what follows, we identify potential sources of systematics and detail our tests to constrain them.

#### 1. Large-scale structure

We expect this item to be one of the main sources of error. We use galaxy+$N$-body simulations to access the sample variance uncertainty in the spectroscopic redshift distribution of the training set. The photometry of the simulation assumed substantially deeper observations, so the selection given below is only roughly appropriate. We first made a cut on $r < 21.8$ followed by a cut on $u < 24.7$ to match the fraction of objects removed by the $u$-band selection on the real data. For simplicity, we only simulate the magnitude-limited surveys of the training set. In addition, because of the overlap between zCOSMOS and one of the PRIMUS fields, we neglect the zCOSMOS sample in the error estimation to simplify the calculation. This approach results in a $\sim 10\%$ increase in the error bars relative to including zCOSMOS as an independent sample. The predicted error bars are overlaid on the simulated overall redshift distribution in Figure 7, and the values of the errors are given in Table 2. The uncertainty in the training-set redshift distributions is not identical to that of the uncertainty in the estimated redshift distributions $N(z)$ derived using the weights, so the error bars should be thought of as approximate. A more detailed estimation of the errors would require SDSS-specific photometry+$N$-body simulations. Relative to the error bars in the training set, the error bars in the weighted $N(z)$ should be (very roughly) about 10%–30% smaller.

---

23 http://www.sdss3.org/dr8/data_access.php#VAC

24 Simulations provided courtesy of Risa Wechsler and Michael Busha. See M. Busha et al. (2012, in preparation) for details.
with increased anti-correlations between neighboring bins, but a more exact statement would require a significantly more detailed investigation. We explore these issues in more detail, and for a different data set, in Cunha et al. (2012).

2. Selection in non-observables. Two of the surveys comprising our training set have selections in observables that are not included in the SDSS magnitude-limited sample. As mentioned previously, the DEEP2-nonEGS sample is selected using $BRI$ photometry to target galaxies above $z > 0.7$. As shown in Cunha et al. (2009), the use of DEEP2 in earlier versions of this catalog resulted in a bump in the overall estimated redshift distribution around $z \sim 0.8$. The present data release has a brighter magnitude cut and additional training data, which has eliminated this bias. DEEP2-nonEGS carries about 1.4% of the total weight. The 2SLAQ sample targets LRGs. Besides SDSS magnitudes, 2SLAQ also uses morphological information in the selection. Because shape correlates poorly with redshift, biases due to inclusion of the 2SLAQ sample are expected to be small. 2SLAQ is an important part of our sample because it provides a better training set for LRGs at higher redshift than the SDSS sample.

3. Spectroscopic redshift failures. The impact of spectroscopic failures is the most difficult to quantify. We chose a bright $r$-magnitude cut and relatively stringent cuts on spectroscopic quality in order to minimize the effect of incorrect redshifts. However, this does increase the incompleteness of the sample. We estimate a completeness rate of roughly 70% for the main training sets comprising our sample (PRIMUS, zCOSMOS, VVDS, and DEEP2-EGS). Here, we define the completeness as the number of objects with high-redshift confidence divided by total number of galaxies targeted for observation. The incompleteness in the training sample is problematic if it is not purely random and if it is not well described by the observables. Nakajima et al. (2012) conducted tests on the selection of the main samples comprising our training set, namely, PRIMUS, VVDS, DEEP2, and zCOSMOS, and found that, with the exception of VVDS, none of the samples showed signs of incompleteness in one-dimensional slices through color and magnitude space.25

In addition, incompleteness that is well localized in the observable space is partly corrected by the weighting procedure. The risk remains that the incompleteness is partly associated with galaxies that are localized in redshift but not in color, or are localized in color, but only comprise a very biased sample of the redshift distribution of galaxies with the same color. A proper assessment of the exact nature of the incompleteness in the training sets would require detailed simulations of the spectroscopic surveys. It is encouraging that the redshift distributions of the different training samples are roughly consistent with each other, despite being obtained from surveys with significantly different instruments and redshift success rates (Figure 7 of Nakajima et al. 2012). If all of the surveys that were used to construct the training sample were unable to obtain redshifts for some specific type of galaxy, then this comparison between their redshift distributions could not be used to diagnose incompleteness issues. However, given the relatively high completeness for some of those samples and the relatively low-redshift, bright sample we are using, this scenario seems unlikely. The completeness could be increased by adopting less stringent quality cuts, at the cost of increasing the fraction of wrong redshifts. C. E. Cunha et al. (2012, in preparation) show, using spectroscopic/photometric simulations of the Dark Energy Survey and spectroscopic follow-up surveys, that, whereas incompleteness in the spectroscopic samples can be robustly identified with colors, incorrect redshifts need to be exquisitely controlled. We therefore, prefer to adopt more stringent quality cuts.

4. Seeing. Nakajima et al. (2012) report that differences between the seeing distribution of the galaxies in the photometric and the training set can lead to biases in the photo-$z$ error calibration. In Figure 8, we show the seeing distributions for all of our photometric sample compared with the four highest weight training samples, not including SDSS, for which the seeing distribution is a near perfect match. The distributions are qualitatively similar, but with a trend to better seeing for the training-set matches. More quantitatively, we checked the sensitivity of our results to seeing-induced biases by including seeing as a variable in the weights estimation. We find only negligible change in the recovered redshift distribution. Hence, although differences in seeing are in general a concern, we find little effect in our data.

Lima et al. (2008) consider a case that contains much of the observational issues described above. In Figure 9 of that paper we can see the effect of using the weights technique to reconstruct the redshift distribution of the DEEP2-EGS sample using a combination of spectroscopic samples. In the case shown, sample variance in the DEEP2-EGS sample is a main limiting factor to the quality of the reconstruction.

For individual $P(z)$’s, the main source of uncertainty is shot noise, because only 100 galaxies were used to estimate each $P(z)$. The choice to fix the number of neighbors keeps the shot

25 The sample in that paper was not purely flux limited, as the requirement that galaxies be well-resolved eliminates a fraction of galaxies that are a weak function of magnitude for $r < 21.5$ and a strong one for $21.5 < r < 21.8$. Thus, the mean magnitude of that sample is brighter than the one in this paper by $\sim 0.1$ mag.
noise equal for all galaxies, but can yield biases or an artificial broadening of the $P(z)$ if the training set is too sparse near the galaxy of interest. However, we do not find the volume spanned by the 100 nearest neighbors to be a good indicator of the $P(z)$ quality, because other properties of the redshift-observable hyper-surface affect the local density of galaxies. A potentially more interesting indicator of bias in individual $P(z)$'s is the distribution of observed properties for the training-set nearest neighbors relative to the galaxy for which a $P(z)$ is needed, i.e., there could be a bias if the galaxy is very offset from the center of the distribution of neighbors. We leave these explorations for a future work.

9. PROPER USE

In this section, we describe the proper use of these redshift distributions. We risk an overly pedantic discussion in order to ensure that past mistakes in these types of analyses are not repeated.

If one desires to use the $P(z)$ to evaluate any nonlinear function $F(z)$, one must integrate the function times the $P(z)$ over the entire distribution; i.e., one must take the expectation value of the function. The reason is quite simple. In general, a function evaluated at the expectation value of a variable is quite different from the expected value of the function. The expectation value of a function should be computed as $\langle F(z) \rangle = \int_0^\infty F(z)P(z)dz$.

It is not correct to simply take the effective redshift $\int zP(z)dz$ and evaluate the function at that redshift.

This statement is true in most interesting science cases. An excellent example is in gravitational lensing, where one must estimate the “critical surface density” $\Sigma_{\text{crit}}$, which determines the lensing strength of a given lens–source pair; the lensing deflection angle is proportional to $\Sigma_{\text{crit}}$. The function $\Sigma_{\text{crit}}$ depends on the angular diameter distances to the lens, source and between lens and source in a nonlinear manner. The proper estimator for a lens at redshift $z_l$ and source with $P(z_s)$ is

$$\Sigma_{\text{crit}}^{-1}(z_l, z_s) = \int_0^\infty \Sigma_{\text{crit}}^{-1}(z_l, z_s)P(z_s)dz_s.$$  \hspace{1cm} (11)

10. $P(z)$ AND GALAXY–GALAXY LENSING: PROOF OF PRINCIPLE

The sensitivity of observational methods to the properties of the $P(z)$ or $N(z)$ depends on the details of how the observation and analysis are performed. In this section, we use the galaxy–galaxy lensing calibration method from Mandelbaum et al. (2008) and Nakajima et al. (2012) as an example of determining this sensitivity. This methodology requires the use of a fair subsample of source galaxies with spectroscopic redshifts. For the purpose of this paper, we use the DEEP2 EGS region, in which there are 730 galaxies that (1) pass all cuts to be included in the SDSS source catalog from Mandelbaum et al. (2005), (2) have secure redshifts from DEEP2, and (3) pass the additional cut $r < 21.5$. DEEP2 EGS is only one of the many training samples used in our analysis, so this exercise should be thought of as a proof of principle.

In brief, we have measured the expected calibration bias $b_z$ on the galaxy–galaxy lensing signal due to the method of estimating the source redshift (i.e., a multiplicative systematic error). This $b_z$ tells us about systematic errors in our conversion from the observed weak lensing shear (or shape distortion) to surface mass density. Quantitatively, the ratio of the true to the estimated surface mass density is $1 + b_z$. We also estimate the degree to which the variance in the lensing signal deviates from the ideal variance we would achieve with optimal weighting by the true source redshift (large deviation results in increased statistical error). The increase in statistical error when we have degraded redshift information arises both from source misidentification, and also from deviations of the weights from the optimal $1/\Sigma_{\text{crit}}^2$. Schematically, these two quantities can be determined via weighted sums over lens–source pairs $j$ (with weight $\tilde{w}_j$), in what follows, estimated quantities using approximate redshift information have a tilde, and ones that use the true redshift do not:

$$b_z + 1 = \frac{\sum_j \tilde{w}_j (\tilde{\Sigma}_{\text{crit}}^{-1}/\Sigma_{\text{crit}}^{-1})_j}{\sum_j \tilde{w}_j} \hspace{1cm} (12)$$

and

$$\text{Variance ratio} = \frac{\text{Ideal variance}}{\text{Real variance}} = \frac{\left(\sum_j \sqrt{\tilde{w}_j/w_j}\right)^2}{\left(\sum_j w_j\right)} \left(\sum_j \frac{1}{\tilde{w}_j}\right) \hspace{1cm} (13)$$

For more detail, see the aforementioned papers.

In Figure 9, we show the results of these calculations for several test cases. First, the red short-dashed curve provides, as a baseline, the calibration bias (top) and variance ratio (bottom) when using the ZEBRA photo-$z$ studied in Nakajima et al. (2012). As shown, there is a significant bias in the lensing signal that must be calibrated. Next, the green long-dashed line shows what happens if we use the $N(z)_{\text{real}}$ as an estimate of the redshift distribution, rather than using any individual galaxy photo-$z$ or $P(z)$ information. Crucially, the lensing signal is unbiased in this case. However, as shown in the bottom panel, we do find an increased statistical error due to lack of redshift information on a per-galaxy basis.

Third, the solid black line demonstrates what happens when we use the individual $P(z)$’s to estimate $\Sigma_{\text{crit}}$ using Equation (11). These $P(z)$’s are derived from a very specific, idealized case, using only EGS both as the training sample and the photometric sample. In this case, the individual $P(z)$’s are on average 40% broader than the DR8 $P(z)$’s because of the use of 100 neighbors to construct each $P(z)$ when the training sample itself is only seven times as large. To compensate for the bias introduced by the small size of the training sample, we have imposed a multiplicative correction factor to the $P(z)$’s such that $\sum P(z) = N(z)_{\text{real}}$ using Equation (4). Nonetheless, there is a calibration bias due to the very significant width of the $P(z)$’s (which can be removed using a calibration sample); but the variance ratio is still far closer to optimal than when we did not use weighting information, and slightly closer than when we used ZEBRA photo-$z$.

The magenta dot-dashed line shows the results when seven neighbors are used to estimate the $P(z)$, not including the galaxy itself. The blue dot-long-dashed line shows the same case but

---

26 Optimal weighting would also include a factor that downweights galaxies with noisier shape measurements, $\propto (r_{\text{mes}}^2 + \sigma_i^2)^{-1}$. For simplicity, we neglect this factor in the tests that follow; however, in order to use this weighting, which modifies the effective $N(z)$, the shape measurement error weighting must also be used in the derivation of the $P(z)$ from the training sample.
with the Equation (4) correction. This use of seven neighbors reduces the abnormally broad \( P(z) \)'s caused by using such a small training sample and 100 neighbors. The mean \( P(z) \) width for the seven neighbors case is 0.0989, to be compared to the mean width of the DR8 \( P(z) \)’s of 0.0983. The calibration bias for seven neighbors is also quite close to the ideal case with \( N(z)_{\text{wei}} \), and the weighting is the closest to optimal of all the cases considered in this paper.

To summarize, we have demonstrated for this simplified training set that, for the purpose of lensing, we achieve a near-perfect signal calibration when using \( N(z)_{\text{wei}} \); i.e., no individual galaxy redshift information. However, the weighting is suboptimal. When we use individual \( P(z) \), the lensing signal can be biased due to their finite width even if \( \sum P(z) = N(z)_{\text{wei}} \), but this bias can be calibrated. The advantage of using individual \( P(z) \) information is that statistical errors on the lensing signal are reduced due to more optimal weighting. This is because a signal-to-noise ratio weighting is proportional to \( \Sigma^{-2} \), so sources expected to be behind the lens are given higher weight than those expected to be close to or in front of the lens.

Again, we emphasize that this analysis used only DEEP2 EGS, and should be used as a proof of principle to gain intuition. Further tests using real data can be found in, e.g., Lima et al. (2008) and Carnero et al. (2012). For tests using simulations, see Cunha et al. (2009). Users of these data should perform similar analyses to these but matched to their exact analysis and selection criteria.

11. SUMMARY

In this paper, we presented a catalog of photometric redshift probability distributions for the SDSS DR8. With some modifications, our method is the same as that used to generate the \( P(z) \) catalog for SDSS DR7, presented in Cunha et al. (2009). For this catalog, we used the ubercal photometry (Padmanabhan et al. 2008). We also included the PRIMUS galaxy sample, which more than doubles the number of galaxies in our training set that are drawn from a flux-limited sample other than SDSS. The addition of PRIMUS provided a significant increase in the total area of the non-SDSS training set, which reduces the sample variance. We examined several potential sources of error, including shot noise, sample variance, seeing, star–galaxy separation, and spectroscopic failures. We expect that sample variance is the main source of uncertainty in our overall redshift distribution. For individual \( P(z) \)’s, shot noise is the limiting uncertainty, since each \( P(z) \) is based on 100 training-set galaxies. These \( P(z) \)’s, and the ensemble \( N(z) \) derived in this work (Table 2), should be useful for a variety of science applications, such as galaxy angular two-point correlation functions, galaxy cluster detection and weak gravitational lensing.

E.S. is supported by DOE grant DE-AC02-98CH10886. C.C. was supported by DOE OJI grant under contract DEFG02-95ER40899 and the Kavli Fellowship at Stanford.

Thanks to Don Schneider for a careful reading of the manuscript and many helpful suggestions. Thanks to the anonymous referee whose excellent comments led to significant improvement of this paper.

Funding for the DEEP2 survey has been provided by NSF grants AST95-09298, AST-0071048, AST-0071198, AST-0507428, and AST-0507483 as well as NASA LTSA grant NNG04GC89G.

We thank the PRIMUS team for sharing their redshift catalog. Funding for PRIMUS has been provided by NSF grants AST-0607701, 0908246, 0908442, 0908354, and NASA grant 08-ADP08-0019. This paper includes data gathered with the 6.5 m Magellan Telescopes located at Las Campanas Observatory, Chile.

A portion of the data presented herein were obtained at the W. M. Keck Observatory, which is operated as a scientific partnership among the California Institute of Technology, the University of California, and the National Aeronautics and Space Administration. The Observatory was made possible by the generous financial support of the W. M. Keck Foundation. The DEEP2 team and Keck Observatory acknowledge the very significant cultural role and reverence that the summit of Mauna Kea has always had within the indigenous Hawaiian community and appreciate the opportunity to conduct observations from this mountain. Funding for the SDSS and SDSS-II has been provided by the Alfred P. Sloan Foundation, the Participating Institutions, the National Science Foundation, the U.S. Department of Energy, the National Aeronautics and Space Administration, the Japanese Monbukagakusho, the Max Planck Society, and the Higher Education Funding Council for England. The SDSS Web site is http://www.sdss.org/.

The SDSS is managed by the Astrophysical Research Consortium for the Participating Institutions. The Participating Institutions are the American Museum of Natural History, Astrophysical Institute Potsdam, University of Basel, University of Cambridge, Case Western Reserve University, University of Chicago, Drexel University, Fermilab, the Institute for Advanced Study, the Japan Participation Group, Johns Hopkins University, the Joint Institute for Nuclear Astrophysics, the Kavli Institute for Particle Astrophysics and Cosmology, the Korean Scientist Group, the Chinese Academy of Sciences (LAMOST), Los Alamos National Laboratory, the
Max-Planck-Institute for Astronomy (MPIA), the Max-Planck-Institute for Astrophysics (MPA), New Mexico State University, Ohio State University, University of Pittsburgh, University of Portsmouth, Princeton University, the United States Naval Observatory, and the University of Washington.

REFERENCES
Abazajian, K. N., Adelman-McCarthy, J. K., Agüeros, M. A., et al. 2009, ApJS, 182, 543
Abrahamse, A., Knox, L., Schmidt, S., et al. 2011, ApJ, 734, 36
Adelman-McCarthy, J. K., Agüeros, M. A., Allam, S. S., et al. 2006, ApJS, 162, 38
Aihara, H., Allende Prieto, C., An, D., et al. 2011, ApJ, 734, 11
Aihara, H., Allende Prieto, C., An, D., et al. 2011, ApJS, 193, 29
Arnouts, S., Cristiani, S., Moscardini, L., et al. 1999, MNRAS, 310, 540
Baum, W. A. 1962, in IAU Symp. 15, Problems of Extra-Galactic Research, ed. G. C. McVittie (Cambridge: Cambridge Univ. Press), 390
Blandford, D. J., Strauss, M. A., et al. 2005, AJ, 129, 2562
Bordoloi, R., Lilly, S. J., & Amara, A. 2010, MNRAS, 406, 881
Cannon, R., Drinkwater, M., Edge, A., et al. 2006, MNRAS, 372, 425
Carnero, A., Sanchez, E., Crocce, M., Cabre, A., & Gaztanaga, E. 2012, MNRAS, 419, 1689
Coe, D., Benitez, N., Sanchez, S. F., et al. 2006, AJ, 132, 926
Coil, A. L., Blanton, M. R., Burles, S. M., et al. 2011, ApJ, 741, 8
Connolly, A. J., Csalai, I., Szalay, A. S., et al. 1995, AJ, 110, 2655
Crocce, M., Gaztanaga, E., Cabre, A., Carnero, A., & Sanchez, E. 2011, MNRAS, 417, 2577
Cunha, C. E., Huterer, D., Busha, M. T., & Wechsler, R. H. 2012, MNRAS, 423, 2379
Cunha, C. E., Lima, M., Oyaizu, H., Frieman, J., & Lin, H. 2009, MNRAS, 396, 565
Eisenstein, D. J., Annis, J., Gunn, J. E., et al. 2001, AJ, 122, 2267
Eisenstein, D. J., Weinberg, D. H., Agol, E., et al. 2011, AJ, 142, 72
Feldmann, R., Carollo, C. M., Porciani, C., et al. 2006, MNRAS, 372, 565
Fukugita, M., Ichikawa, T., Gunn, J. E., et al. 1996, AJ, 111, 1748
Garilli, B., Le Fèvre, O., Guzzo, L., et al. 2008, A&A, 486, 683
Gerdes, D. W., Sypniewski, A. J., McKay, T. A., et al. 2010, ApJ, 715, 823
Gunn, J. E., Carr, M., Rockosi, C., et al. 1998, AJ, 116, 3040
Gunn, J. E., Siegmund, W. A., Mannery, E. J., et al. 2006, AJ, 131, 2332
Hogg, E., Fabricius, C., Makarov, V. V., et al. 2000, A&A, 355, L27
Hogg, D. W., Finkbeiner, D. P., Schlegel, D. J., & Gunn, J. E. 2001, AJ, 122, 2129
Ilbert, O.,Arnouts, S., McCracken, H. J., et al. 2006, A&A, 457, 841
Koo, D. C. 1985, AJ, 90, 418
Lilly, S. J., Le Fèvre, O., Crampton, D., Hammer, F., & Tresse, L. 1995, ApJ, 455, 50
Lilly, S. J., Le Fèvre, O., Renzini, A., et al. 2007, ApJS, 172, 70
Lima, M., Cunha, C. E., Oyaizu, H., et al. 2008, MNRAS, 390, 118
Loh, E. D., & Spillar, E. J. 1986, ApJ, 303, 154
Lupton, R. H., Gunn, J. E., Ivezić, Z., et al. 2001, in ASP Conf. Ser. 238, Astronomical Data Analysis Software and Systems X, ed. F. R. Hamden, Jr., F. A. Primini, & H. E. Payne (San Francisco, CA: ASP), 269
Mandelbaum, R., Hirata, C. M., Seljak, Uroš, et al. 2005, MNRAS, 361, 1287
Mandelbaum, R., Seljak, U., Hirata, C. M., et al. 2008, MNRAS, 386, 781
Nakajima, R., Mandelbaum, R., Seljak, U., et al. 2012, MNRAS, 420, 324
Padmanabhan, N., Schlegel, D. J., Finkbeiner, D. P., et al. 2008, ApJ, 674, 1217
Pier, J. R., Schlegel, D. J., Finkbeiner, D. P., et al. 2003, AJ, 125, 1559
Pushchell, J. J., Owen, F. N., & Liang, R. A. 1982, ApJ, 257, L57
Schlegel, D. J., Finkbeiner, D. P., & Davis, M. 1998, ApJ, 500, 525
Scranton, R., Ménard, B., Richards, G. T., et al. 2005, ApJ, 633, 589
Sheldon, E. S., Johnston, D. E., Frieman, J. A., et al. 2004, AJ, 127, 2544
Smith, J. A., Tucker, D. L., Kent, S., et al. 2002, AJ, 123, 2121
Stoughton, C., Lupton, R. H., Bernardi, M., et al. 2002, AJ, 123, 485
Strass, M. A., Weinberg, D. H., Lupton, R. H., et al. 2002, AJ, 124, 1810
Tucker, D. L., Kent, S., Richmond, M. W., et al. 2006, Astron. Nachr., 327, 821
Weiner, B. J., Phillips, A. C., Faber, S. M., et al. 2005, ApJ, 620, 595
Wirth, G. D., Willmer, C. N. A., Amico, P., et al. 2004, AJ, 127, 3121
Wittman, D. 2009, ApJ, 700, L174
Yee, H. K. C., Morris, S. L., Lin, H., et al. 2000, ApJ, 129, 475
York, D. G., Adelman, J., Anderson, J. E., Jr., et al. 2000, AJ, 120, 1579