CALIBRATING REDSHIFT DISTRIBUTIONS BEYOND SPECTROSCOPIC LIMITS WITH CROSS-CORRELATIONS

Jeffrey A. Newman [1,2]
University of Pittsburgh, Pittsburgh, PA 15260
Accepted: *May 5, 2008*

ABSTRACT

We describe a new method that can measure the true redshift distribution of any set of objects that are studied only photometrically. Measuring the angular cross-correlation between objects in the photometric sample with objects in some spectroscopic sample as a function of the spectroscopic z, along with other, standard correlation measurements, provides sufficient information to reconstruct the redshift distribution of the photometric sample. The spectroscopic sample need not resemble the photometric sample in galaxy properties, but must fall within its sky coverage. We test this hybrid, photometric-spectroscopic cross-correlation technique with Monte Carlo simulations based on realistic error estimates (including sample variance). RMS errors in recovering both the mean redshift and σ of the redshift distribution for a single photometric redshift bin with true distribution given by a Gaussian are $1.4 \times 10^{-3} (\sigma_z/0.1)(\Sigma_p/10)^{-0.3}(dN_s/dz/25,000)^{-1/2}$, where σ_z is the true Gaussian σ, Σ_p is the surface density of the photometric sample in galaxies/arcmin^2, and dN_s/dz is the number of galaxies with a spectroscopic redshift per unit z. We test the impact of non-Gaussian redshift outliers and of systematic errors due to unaccounted-for bias evolution, errors in measuring autocorrelations, photometric zero point variations, or mistaken cosmological assumptions, and find that none will dominate measurement uncertainties in reasonable scenarios. The true redshift distributions of even arbitrarily faint photometric samples may be determined to the precision required by proposed dark energy experiments ($\Delta\langle z \rangle \lesssim 3 \times 10^{-3}$ at $z \sim 1$) with this method.

Subject headings: galaxies: distances and redshifts, cosmology: large-scale structure of universe, methods: miscellaneous, surveys

1. INTRODUCTION

Almost all cosmological tests require information about the distance or redshift of the objects studied. For instance, the comoving length scale corresponding to baryon acoustic oscillations should remain fixed over time, but the corresponding angular size will depend on redshift in a cosmology-dependent manner; hence, if we measure this angular scale as a function of redshift, we may infer cosmological parameters (Seo & Eisenstein 2003). Similarly, measuring weak lensing strength as a function of redshift can provide strong constraints on cosmological models (Kaiser 1998), but the observed weak lensing signal will depend sensitively upon the redshift distribution of the background objects studied (Huterer 2002).

However, although redshift (z) information is required for interpretation, it is infeasible to measure spectroscopic redshifts for the samples of hundreds of millions of extremely faint galaxies to be studied by proposed photometric dark energy probes such as the Large Synoptic Survey Telescope (LSST; Tyson & Angel 2001; Tyson 2005) or the Supernova / Acceleration Probe (SNAP; Deustua et al. 2000; Perlmutter & SNAP 2004), or even the millions of faint galaxies in samples now underway (Hoekstra et al. 2006). Hence, these projects will make use of photometric information to infer redshift distributions and to allow objects to be divided into multiple

redshift bins for analysis. This is possible because galaxy spectra are generally not featureless; as a spectrum redshifts through photometric passbands, its measured colors will vary with z in ways that may be predicted based on spectroscopic observations of galaxies at similar redshifts or from the spectra of local analogues.

However, these "photometric redshifts" in general lack the precision of spectroscopically-derived redshifts, both because of photometric noise and outliers (e.g. in cases of overlapping galaxies) and because in some classes of galaxies (e.g. those forming stars most rapidly) the observable spectral features are weak at most redshifts (see, e.g., Ilbert et al. 2006). The true redshift distribution of objects with a given photometric redshift value may or may not be strongly or even singly-peaked, depending on the galaxy type, passbands used, photometric errors, etc.

Because of these difficulties, dark energy experiments are unlikely to ever treat individual photometric redshifts as known with precision; instead, the approach taken in forecasts is to assume that objects will be divided into photometric redshift bins (Albrecht et al. 2006). However, in order to obtain precision measurements of the properties of dark energy, both weak lensing and photometric baryonic acoustic oscillation (BAO) experiments require that the true redshift distribution of the objects in each bin be known with very high accuracy. Projections for SNAP are that any overall bias in the mean redshift of a bin must be smaller than $2 - 4 \times 10^{-3}$ in z for dark energy constraints not to be degraded strongly (Huterer et al. 2004; Ma et al. 2006; Huterer et al. 2006); for LSST, it is estimated that the

Electronic address: janewman@pitt.edu

[1] Hubble Fellow
[2] present address: Department of Physics and Astronomy, University of Pittsburgh, 3941 O'Hara St., Pittsburgh, PA 15260

mean redshift in each bin must be known to $\sim 2 \times 10^{-3}(1+z)$ (Zhan & Knox 2006; Zhan 2006; Knox et al. 2006; Tyson 2006,Tyson, Connolly, & Newman, in prep.). The true width of each bin must also be known, though with less precision ($\Delta\sigma_z < \sim 3 \times 10^{-3}(1+z)$ for LSST, where σ_z is the Gaussian sigma of the true redshift distribution; Tyson, Connolly, & Newman, in prep.).

These targets will be difficult to meet with standard spectroscopic techniques. Both ongoing and proposed experiments reach depths far too faint (up to $R_{AB} \sim 30$ in the case of SNAP) for existing telescopes and spectrographs to measure redshifts. Recent and ongoing surveys of faint galaxies using the largest telescopes available have obtained spectra for tens of thousands of galaxies to $R \sim 24$ or $I \sim 23$ (Davis et al. 2006, in prep; Le Fèvre et al. 2005) and for a few thousand selected galaxies to $R \sim 25.5$ (Steidel et al. 1999, 2003, 2004), but precise measurements of the redshift distributions of samples of galaxies with $R \sim 26 - 27$ or even fainter will be required in the future if proposed surveys are to reach their targets.

Efforts to obtain true redshift distributions spectroscopically are made more difficult by the fact that faint galaxy surveys fail to obtain redshifts for a substantial fraction of their targets. The DEEP2 Galaxy Redshift Survey, for instance, has obtained secure redshifts for $\sim 70\%$ of the galaxies studied (Cooper et al. 2006). Roughly half of the missed targets appear to be starforming galaxies at $z > 1.4$ (which have no features within the DEEP2 spectral window) based on follow-up observations of blue DEEP2 redshift failures (C. Steidel, priv. comm.), but the redshift distribution of the remainder is unknown and currently being tested.

Surveys of fainter galaxies have even lower success rates. Despite integration times of more than thirty hours per object, the Gemini Deep Deep Survey (GDDS) only succeeded in measuring spectroscopic redshifts for $\sim 15\%$ of their targets with $24 < I < 24.5$, and $< 50\%$ of objects with $23 < I < 24$ (Abraham et al. 2004). However, even with a completeness as high as the Sloan Digital Sky Survey ($\sim 99\%$; Schlegel et al. 2007, in prep.), if the objects missed are not a random subsample, redshift distributions could be biased beyond the tolerances of future dark energy surveys.

Despite these difficulties, it is essential that there be some external method for testing photometric redshifts of faint galaxies, as it is quite likely that the spectral energy distributions (SEDs) of bright galaxies (whose redshifts are more easily obtained) should differ from those for fainter objects. Both locally and at $z \sim 1$, the bluest galaxies (in rest-frame color) are intrinsically faint; they have no luminous analogues. These issues make using SEDs from bright galaxies to determine photometric redshifts for fainter galaxies problematic. Furthermore, at fainter magnitudes, higher-redshift galaxies will be more and more prevalent within a given photometric redshift bin. These high-redshift galaxies may have contributions to their SEDs from metal-deficient "Population III" stars that appear to have no local analogues (Jimenez & Haiman 2006). As an additional complication, Population III contributions should be greater in fainter, lower-mass galaxies than in more massive galaxies at all redshifts, given the evidence that lower-mass objects generally start forming stars later (e.g., Noeske et al., submitted).

Despite these difficulties, if photometric dark energy surveys are to reach their goals, it is vital that we have some method of calibrating photometric redshifts with high precision (Albrecht et al. 2006). In this paper, we describe a new method that can determine the true redshift distribution for any class of object (e.g. objects in a particular photometric redshift bin) by exploiting the fact that all galaxies at a given redshift cluster with each other. We presume the existence of a large sample (or samples) of objects with spectroscopic redshifts. The observed angular clustering between any two samples of galaxies will depend on both the intrinsic clustering of objects in the two samples with each other, and the degree to which they overlap in redshift (since clustering over extremely large distances is minimal). Hence, by measuring the apparent angular cross-correlation between the positions of the photometric objects and the spectroscopic sample *as a function of the spectroscopic z*, we may determine the actual redshift distribution of objects in the unknown class; the information provided by autocorrelation measurements for each sample allows us to break the degeneracy between correlation strength and redshift distribution.

Similar cross-correlation techniques have been used in the past to measure correlation functions (Phillipps 1985; Masjedi et al. 2006) and luminosity functions (Phillipps & Shanks 1987) at separations or depths where redshift surveys are incomplete; here, we explore their use to measure redshift distributions. The angular cross-correlation can be measured with good precision in uniform, well-calibrated photometry, as required for future dark energy probes, as there will be many photometric galaxies near each spectroscopic galaxy on the sky. The use of angular cross-correlations between photometric redshift bins to constrain the presence of redshift outliers has also been explored (Schneider et al. 2006; Padmanabhan et al. 2006), but cannot determine redshift distributions in detail.

The principal requirement of this method is that redshift survey data be available overlapping the photometric sample; however, the objects with redshifts need not be similar to the target class (e.g., only high-confidence redshifts of relatively bright galaxies could be used when determining the redshift distribution of a sample of very faint galaxies). In §2, we provide the theoretical underpinnings of this method. In §3 we present Monte Carlo tests of its effectiveness for scenarios appropriate for current and future redshift surveys. We evaluate potential sources of systematic error in §4, and in §5 we conclude. Throughout this paper, we will use comoving coordinates for all distances and assume a cosmology with zero spatial curvature. Where a specific cosmology must be adopted, we assume a flat ΛCDM cosmology with matter density $\Omega_m = 0.3$, dark energy density $\Omega_\Lambda = 0.7$, and Hubble parameter $H_0 = 100h$ km s^{-1} Mpc^{-1}.

2. MEASURING REDSHIFT DISTRIBUTIONS VIA CROSS-CORRELATIONS

2.1. *Basic Techniques*

Consider two sets of objects at cosmological distances, one with secure redshift measurements, which we will label 's' (for 'spectroscopic', though exceedingly accurate

photometric redshifts might be used), and the other with unknown redshifts, which we will label 'p' (for 'photometric'). In the most likely applications, the photometric sample would be a subset of objects in some photometric dataset – e.g. objects in some bin of photometric redshift – and the spectroscopic sample would result from one or more redshift surveys within the region of sky covered by p. Although cross-correlation analyses require significant sample sizes, it should be possible to determine the bias and uncertainty in photometric redshifts as a function of z by studying samples in a set of photometric-redshift bins.

The mean comoving number density of objects in the photometric sample (p) a comoving real-space distance r from an object in the spectroscopic sample (s) at redshift z, $n_p(r, z)$ will be

$$\langle n_p(r, z) \rangle = n_p(z)(1 + \xi_{sp}(r, z)), \qquad (1)$$

where $n_p(z)$ is the comoving number density of objects in sample p at redshift z and $\xi_{sp}(r, z)$ is the two-point cross-correlation function between samples s and p. Hence, ξ_{sp} defines the excess probability of finding an object of class p separated by a distance r from an object of class s that is at redshift z, above the probability if the two populations do not cluster with each other. Finally, we denote the probability distribution function for the true redshift of an object in the photometric sample by $\phi_p(z)$.

We assume that the surveyed objects are distant from us compared to the length over which correlations are significant and that $\xi_{sp}(r, z)$, $n_p(z)$, and $\phi_p(z)$ may all be treated as constant over separations in the redshift direction comparable to that length (assumptions that all hold in typical high-redshift samples). In the distant-observer approximation, we may define $r^2 = \pi_l^2 + r_p^2 = \pi_l^2 + d_A(z)^2\theta^2$, where π_l is the comoving separation between two objects along the line-of-sight direction, r_p is their projected comoving separation in the plane of the sky, $d_A(z)$ is the angular size distance to redshift z, and θ is their angular separation in radians. We also define $l(z)$ to be the comoving distance to redshift z given by $l(z) = \int_0^z c/H(z)dz$, where c is the speed of light and $H(z)$ the Hubble expansion parameter at redshift z. In calculations for measurements of angular correlations, we may ignore redshift-space distortions, so for a photometric object at redshift z' separated by π_l along the line-of-sight direction from a spectroscopic object at redshift z, $l(z') = l(z) + \pi_l$.

The fundamental quantity we wish to recover is $\phi_p(z)$, the probability distribution for the true redshift of an object in p. It can be related to n_p and the mean surface density of objects in p on the sky (in units of objects per steradian), denoted here by Σ_p:

$$\phi_p(z) = \frac{dN_p}{dz\,d\Omega} \bigg/ \left(\int_0^\infty \frac{dN_p}{dz'\,d\Omega}\,dz' \right)$$
$$= n_p(z)\frac{dV}{dz\,d\Omega} \bigg/ \left(\int_0^\infty n_p(z')\frac{dV}{dz'\,d\Omega}\,dz' \right)$$
$$= n_p(z)\frac{dV}{dz\,d\Omega} \bigg/ \Sigma_p$$
$$= \frac{n_p(z)}{\Sigma_p}d_A(z)^2\frac{dl}{dz}, \qquad (2)$$

where $dN_p/(dz\,d\Omega)$ is the number of objects in the pho-

tometric sample per unit redshift per steradian and $dV/(dz\,d\Omega)$ is the amount of comoving volume per unit redshift per steradian (equal to $d_A(z)^2\,dl/dz$).

Equation 1 gives the excess number density of photometric objects near a spectroscopic object as a function of their real-space separation and the spectroscopic object's redshift. However, what we are actually able to measure is the excess number of objects per unit area on the sky. We therefore multiply Equation 1 by $dV/dz\,d\Omega$ and integrate over the possible redshifts of a photometric object, z', to obtain $\langle\Sigma(\theta, z)\rangle$, the mean surface density of objects in p an angle θ from an object in s at redshift z. We then obtain:

$$\langle\Sigma(\theta, z)\rangle = \int_0^\infty n_p(z')d_A(z')^2\frac{dl}{dz'}\,dz'$$
$$+ \int_0^\infty \xi_{sp}(r, z)n_p(z')d_A(z')^2\frac{dl}{dz'}\,dz'$$
$$= \int_0^\infty \Sigma_p\,\phi_p(z')\,dz' + \int_0^\infty \xi_{sp}(r, z)\,\Sigma_p\,\phi_p(z')\,dz'$$
$$= \Sigma_p\,(1 + w_{sp}(\theta, z)), \qquad (3)$$

where $w_{sp}(\theta, z)$ defines the angular cross-correlation function between the spectroscopic and photometric samples and Σ_p is the mean surface density of objects in p over the sky. This constitutes the principal observable we will use to reconstruct the redshift distribution of the photometric sample (sample p).

For convenience, we assume that all correlation functions may be described by power laws with linear biasing. This assumption is somewhat unrealistic – in real applications, precision measurements should use a halo model (cf. Cooray & Sheth 2002 and references therein) or other more sophisticated methods – but is sufficiently accurate to predict uncertainties for cross-correlation methods. If ξ_{sp} is represented by the power law form $\xi_{sp}(r) = (r/r_{0,sp})^{-\gamma}$, then the integrals in Equation 3b may be evaluated analytically to obtain:

$$w_{sp}(\theta, z) = \frac{\phi_p(z)\,H(\gamma)\,r_{0,sp}^\gamma\,\theta^{1-\gamma}\,d_A(z)^{1-\gamma}}{dl/dz}, \quad (4)$$

where $H(\gamma) = \Gamma(1/2)\Gamma((\gamma-1)/2)/\Gamma(\gamma/2)$ (Peebles 1980), and we have treated $\phi_p(z)$ and $d_A(z)$ as constant over the range in z' for which ξ_{sp} is nonnegligible (i.e., where $l(z) - l(z')$ is not much greater than $r_{0,sp}$).

Although future applications of this method may make use of angular information (e.g. to constrain biasing models), when determining the uncertainties resulting from application of cross-correlation methods below, we will focus on the integral of w_{sp} within an angle equivalent to some comoving distance r_{max} (which we will leave fixed with z) for simplicity. We label this integrated w_{sp} $\tilde{w}(z)$, and define the angle corresponding to r_{max} at a given z to be $\theta_{max}(z)$. Integrating Equation 4 over θ, we may relate ϕ_p to $\tilde{w}$ by the equation:

$$\phi_p(z) = \tilde{w}(z)\frac{3-\gamma}{2\pi}\frac{d_A(z)^2\,dl/dz}{H(\gamma)\,r_{0,sp}^\gamma\,r_{max}^{3-\gamma}}. \qquad (5)$$

In general, a useful choice of r_{max} should be large enough that the effects of nonlinear biasing (which can complicate modeling of correlation functions) are small,

but small compared to the angular size of the photometric sample to minimize edge effects. Our fiducial scenario will use $r_{max} = 10\,h^{-1}$ Mpc, corresponding to roughly a quarter of a degree at $z = 1$. Correlation functions at both $z \sim 0$ and $z \sim 1$ are closely approximated by power laws at this scale (Zehavi et al. 2005; Coil et al. 2006c).

Inspection of Equations 4 and 5 shows that to determine $\phi_p(z)$ from w_{sp} or $\tilde{w}$, we must know the basic cosmology (sufficient to determine $d_A(z)$ and dl/dz modulo factors of h), $r_{0,sp}$, and γ. We test the degree to which the cosmology must be known in §4.4; realistic uncertainties in cosmological parameters prove to have negligible impact.

The same observations used to measure w_{sp} provide sufficient information to determine $r_{0,sp}$ and γ via the autocorrelation functions of the photometric and spectroscopic samples, ξ_{pp} and ξ_{ss}. This is true because, under our assumption of linear biasing, the cross-correlation $\xi_{sp}(r)$ must be given by the geometric mean of the autocorrelations of the two samples, $\xi_{sp} = (\xi_{ss} \times \xi_{pp})^{1/2}$; we will refer to this as the "simple biasing" assumption hereafter. This equation holds to high accuracy for the measured cross-correlations between subsamples in modern redshift surveys, even when their clustering differs strongly (Coil et al. 2007). The autocorrelation function of the spectroscopic sample, ξ_{ss}, is measurable directly from the spectroscopic sample, and is in fact a prime observable of redshift surveys; it thus remains only to determine ξ_{pp}.

We may use the observed angular autocorrelation of the photometric sample, $w_{pp}(\theta)$, in conjunction with an initial guess for $\phi_p(z)$ to obtain the mean parameters of ξ_{pp}, since they are related by Limber's equation (evaluated for a power law correlation function, with scale length $r_{0,p}$ a function of z but exponent γ_p constant):

$$ w_{pp}(\theta) = H(\gamma_p)\,\theta^{1-\gamma_p} \int_0^\infty \phi_p^2(z)\, r_{0,p}(z)^{\gamma_p} \, \frac{d_A(z)^{1-\gamma_p}}{dl/dz} \, dz $$

(6)

(Peebles 1980). Note that γ_p can be measured directly from the shape of $w_{pp}(\theta)$, so given a form for ϕ_p and the cosmology, the mean value of $r_{0,p}$ may be determined directly from its amplitude (in general, γ varies only modestly – $< 10\%$ – even amongst samples of galaxies with very different biasing; Zehavi et al. 2005). This procedure may be iterated by using the derived parameters of ξ_{pp} to determine ξ_{sp} and hence ϕ_p from cross-correlations, then redetermining ξ_{pp} using this ϕ_p, and then refining ϕ_p from cross-correlations, etc. until convergence is reached. Although w_{pp} yields only a weighted mean of the value of $r_{0,p}(z)$, the consequences of feasible amounts of variation in the bias of the photometric sample with redshift are modest; we demonstrate this in §4.1.

To summarize: the most basic large-scale structure measurements possible where a photometric sample overlaps a spectroscopic one – the two-point autocorrelation functions of each sample with itself, and their cross-correlation on the sky, measured as a function of spectroscopic redshift – provide sufficient information to reconstruct the redshift distribution of the photometric sample. In the remainder of this paper, we will attempt to determine the uncertainties, both random and system-

atic, that should result from applying such methods to realistic samples.

2.2. Error Estimates

We begin by estimating the error in a measurement of $\phi_p(z)$ in some small bin of redshift of centered at z and of width Δz. We presume here that the errors in ϕ_p will be dominated by the uncertainty due to counting statistics in a measurement of $\tilde{w}$, the integral of w_{sp} within the angle $\theta_{max}(z)$. Poisson uncertainties should dominate when $\tilde{w}$ is small (the "weak-clustering" limit), which should always be the case unless ϕ_p is unrealistically narrow (Peebles 1980). We defer investigation of possible systematic errors to §4.

Modulo the modest impact of sample variance (see §3.4), we expect uncertainties to be dominated by errors in $\tilde{w}$, as the autocorrelations ξ_{ss} and w_{pp} should be measured more precisely. Then, applying standard propagation of errors to Equation 5, $\sigma(\phi_p(z))/\phi_p(z) = \sigma(\tilde{w})/\tilde{w}$. Furthermore, $\sigma(\tilde{w})/\tilde{w}$ must equal the uncertainty in the total excess (over random) number of neighbors in p surrounding any member of s due to clustering (which we will label $\sigma(N_c)$), divided by the expected number of these neighbors (denoted by N_c), as $\tilde{w}$ is directly proportional to N_c by definition. Since these quantities are simple to predict, we will determine $\sigma(\phi_p(z))/\phi_p(z)$ by calculating the equivalent quantity, $\sigma(N_c)/N_c$.

In the weak-clustering limit (i.e., so long as $\tilde{w}$ is small, as is true here), the uncertainty in N_c, $\sigma(N_c)$, is given simply by the Poisson uncertainty in the expected total number of spectroscopic-photometric pairs if there is no clustering (Peebles 1980). Thus,

$$ \sigma^2(N_c) = \left(\Sigma_p\, \pi \theta_{max}^2 \right) \left(\frac{dN_s}{dz} \Delta z \right) $$

$$ = \pi\, \Sigma_p\, \frac{dN_s}{dz}\, \Delta z \left(\frac{r_{max}}{d_A(z)} \right)^2, \qquad (7) $$

where dN_s/dz gives the actual redshift distribution of the spectroscopic sample. Inside the first parentheses in Equation 7 is found the expected number of members of the photometric sample within θ_{max} of each object in the spectroscopic sample, while inside the second parentheses we give the number of objects in s within the designated redshift bin.

We may determine N_c, the total number of excess spectroscopic-photometric $(s-p)$ pairs within separation θ_{max} over random due to correlations, by integrating the real-space two-point cross-correlation function over the relevant volume:

$$ N_c = \left(\int_0^{r_{max}} 2\pi r_p \int_0^\infty n_p(z')\, \xi_{sp}(r,z)\, dz' dr_p \right) \times \left(\frac{dN_s}{dz} \Delta z \right) $$

$$ = \left(\int_0^{r_{max}} 2\pi r_p \int_0^\infty \frac{\phi_p(z')\Sigma_p}{d_A(z')^2\, dl/dz'}\, \xi_{sp}(r,z)\, dz' dr_p \right) $$

$$ \times \left(\frac{dN_s}{dz} \Delta z \right) $$

$$ = 2\pi H(\gamma) \frac{\phi_p(z)\, \Sigma_p}{d_A(z)^2\, dl/dz} r_{0,sp}^\gamma\, \frac{dN_s}{dz}\, \Delta z \int_0^{r_{max}} r_p^{2-\gamma}\, dr_p $$

$$ = \frac{2\pi H(\gamma)}{3-\gamma} \frac{\phi_p(z)\, \Sigma_p}{d_A(z)^2\, dl/dz} r_{0,sp}^\gamma\, r_{max}^{3-\gamma}\, \frac{dN_s}{dz}\, \Delta z \qquad (8) $$

where again we have separated the number of members of the photometric sample around each member of the spectroscopic sample from the total number of members of the spectroscopic sample within Δz using parentheses before combining them, and assumed that $\phi_p(z)$ is approximately constant over the range in $l(z')$ where $\xi_{sp}(r, z)$ is nonngeligible.

Combining Equations 7 and 8, we then find:

$$\sigma(\phi_p(z)) = \phi_p(z)\frac{\sigma(N_c)}{N_c} \qquad (9)$$

$$= \frac{3 - \gamma}{2\sqrt{\pi}H(\gamma)} \left(\Sigma_p \frac{dN_s}{dz} \Delta z\right)^{-1/2} \frac{d_A(z)\, dl/dz}{r_0^\gamma\, r_{max}^{2-\gamma}}.$$

It is worth noting that these errors scale only very slowly with r_{max}, since for typical galaxy samples $\gamma \approx 1.6 - 1.9$. As a consequence, it is possible to minimize nonlinear effects not only by measuring correlations to large separations (i.e., increasing r_{max}), but also by excluding the smallest separations ($r_p < 1 - 2h^{-1}$ Mpc) from the calculation of the integrated correlation, $\tilde{w}(z)$. Predicted measurement uncertainties increase only modestly if small-separation pairs are not considered; for $r_{max} = 10h^{-1}$ Mpc, excluding the central $2h^{-1}$ Mpc increases overall errors by roughly 15%. That exclusion radius is larger than the maximum r_p where non-power law cross-correlations have been observed in studies of the clustering of blue, starforming galaxies about galaxy groups (Coil et al. 2006a), which is likely to be a maximally pathological case.

2.3. Sample Variance

These error estimates have ignored the fact that the mean density of the regions where we perform these cross-correlation measurements at a given z may be higher or lower than the Universal mean, the effect commonly referred to as "sample" or "cosmic" variance. Thus, the recovered redshift distribution of the photometric sample, $\phi_r(z)$, will differ from the distribution that would be obtained if an infinite volume were surveyed, $\phi_p(z)$. If the region over which cross-correlations are measured corresponds to the full area covered by the photometric sample, then in fact ϕ_r may be the desired quantity, rather than the "true", underlying distribution, ϕ_p. Outside of this regime, we must consider the impact that sample variance will have on our recovery of $\phi_p(z)$.

In particular, we can place two limits on the impact of sample variance. If cross-correlations are measured over very large areas of sky (hundreds of square degrees), sample variance should be negligible compared to other sources of error, and the random errors in the recovery of $\phi_p(z)$ will simply be given by Equation 9. We will assess how much area is sufficient to reach this regime in §3.4. Thus, our previous error estimates are in fact a lower limit on measurement uncertainties from cross-correlation techniques.

If only a few fields with small areas are surveyed, the impact of sample variance is much greater. However, the spectroscopic sample may be used to limit this impact, as the variations in density will cause proportional variations in the number of galaxies found in a given redshift bin (compared to a smooth model). Since $w_{sp}(z)$ measures the excess number of companions per spectroscopic

object at a given z, these variations in the redshift distribution of the spectroscopic sample, s, do not affect the measured $\phi_p(z)$ directly. However, there will be corresponding variations in the number of members of the photometric sample p at that redshift, with the amplitude of those variations proportional to the ratio of the large-scale bias of sample p to that of sample s.

Since that ratio of biases would be determined in the process of measuring $\phi_p(z)$ from cross-correlations, we may estimate the universal value of $\phi_p(z)$ from the value reconstructed in a particular region of the sky:

$$\phi_p(z) = \frac{\phi_r(z)}{1 + b_p/b_s\, \Delta_s(z)}, \qquad (10)$$

where b_p and b_s are the linear, large scale biases of the photometric and spectroscopic samples (p and s, respectively) and $\Delta_s(z)$ is the fractional deviation of dN_s/dz at a given redshift from a smooth model; i.e., $[(dN_s/dz)_{observed} - (dN_s/dz)_{true}]/(dN_s/dz_{true})$. Because $(dN_s/dz)_{observed}$ is determined from the finite number of spectroscopic objects within Δz, it will be subject to Poisson variance; hence $\Delta_s(z)$ has a measurement uncertainty $\sigma(\Delta_s) = (dN_s/dz \times \Delta z)^{-1/2}$. This will propagate into a residual uncertainty in $\phi_p(z)$ of $(b_p/b_s)(dN_s/dz \times \Delta z)^{-1/2}\phi_p(z)$ (taking $(1 + (b_p/b_s)\Delta_s)^2 \sim 1$, which holds for all realistic survey characteristics). This error is independent of counting-statistics errors; thus when assessing the maximal impact of sample variance, we combine it with the measurement uncertainty given by given by Equation 9 following standard propagation of errors.

Because it increases overall uncertainties the most where $\phi_p(z)$ is largest, the net effect of sample variance after correcting with $(dN_s/dz)_{observed}$ is to reduce modestly the advantages of samples with tight redshift distributions or high surface densities. Since errors from both counting statistics and sample variance scale as $(dN_s/dz \times \Delta z)^{-1/2}$, though, both of these sources of uncertainty will be reduced by the same fraction if dN_s/dz is increased. In our Monte Carlo simulations, we assume $b_p = b_s$. We expect that for typical datasets $b_p < b_s$, as photometric samples should go fainter than spectroscopic samples, and fainter objects tend to have lower bias (Zehavi et al. 2005; Coil et al. 2006c), making this assumption an upper limit. In the simulations below, we will estimate the errors in the reconstruction of $\phi_p(z)$ both when sample variance is negligible and when dN_s/dz is used for corrections, in order to bracket the possibilities.

3. MONTE CARLO TESTS

3.1. Basic Scenarios

We now investigate the degree to which cross-correlation techniques can recover true redshift distributions for photometric samples. For our most basic scenario, we adopt a simple $\phi_p(z)$ distribution given by a Gaussian with mean redshift z_0 (which we generally take to be 1, near the peak of sensitivity of most dark energy measurement methods) and standard deviation σ_z; i.e.,

$$\phi_p(z) = g(z) = \frac{1}{\sqrt{2\pi}\sigma_z} \exp\left(-\frac{(z - z_0)^2}{2\sigma_z^2}\right). \qquad (11)$$

We then use Monte Carlo techniques to test the recovery of both the mean and standard deviation of this distri-
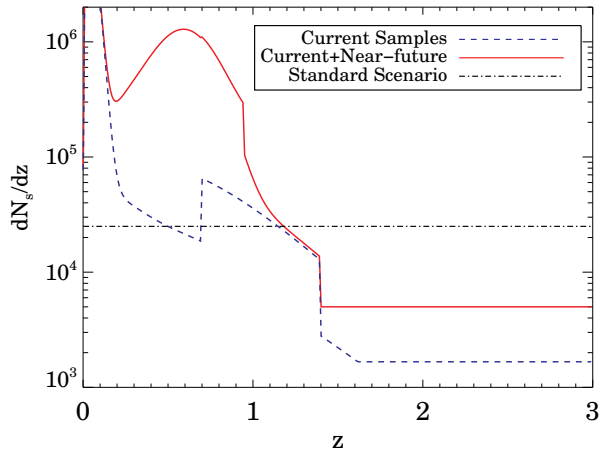
Fig. 1.— Redshift distributions assumed for current spectroscopic samples (blue dashed line) and future samples (red solid line). The assumed characteristics of each sample are given in Table 1. The differences are the addition of an intermediate-redshift survey, PRIMUS (Eisenstein et al. 2007, in prep.); a baryonic oscillation survey, WiggleZ (Glazebrook et al. 2007, in prep.); zCOSMOS (Lilly & The Zcosmos Team. 2005); and larger samples at $z > 2$ in the near-future scenario. These samples were used to produce the Monte Carlo realizations shown in Figure 2. The black, dot-dashed line indicates the assumption used for our standard scaling scenario, which approximates current redshift samples at $z \sim 1$.

bution given a spectroscopic sample with some redshift distribution dN_s/dz. A test of these techniques with catalogs taken from an N-body simulation is now underway, and finding similar results (Wittman 2008, in prep.).

To perform these Monte Carlo tests, we generate realizations of the recovered $\phi_p(z)$ in a large number of bins of width Δz, adding to the true $\phi_p(z)$ in each bin an error drawn randomly from a Gaussian distribution with mean zero and standard deviation given by $\sigma(\phi_p(z))$ for that bin, incorporating both counting statistics and sample variance as described above. Where $\sigma_z > 0.1$, we use bins of width $\Delta z = 0.01$; otherwise, the bin width used is $\Delta z = 0.01 \times (\sigma_z/0.1)$ to ensure the peak is well-resolved. For each realization, we fit for the parameters of ϕ_p with standard nonlinear least-squares techniques. To ensure stability in the fitting, we provide initial guesses for the parameters given by the true value plus a random value drawn from a Gaussian distribution with standard deviation 10% of the true value. This is effectively equivalent to assuming that the true distribution parameters are known to 10%, far worse than the tolerances for most dark energy experiments and much larger than the errors resulting from the cross-correlation measurements. We show example realizations and reconstructions (based on the redshift distributions shown in Figure 1) in Figure 2.

For every scenario tested in this paper, we generate ten thousand realizations of this sort, measuring the parameters of $\phi_p(z)$ each time; we then determine the mean and standard deviation of the results for each parameter to test the efficacy of cross-correlation methods. For our basic scenarios, we vary three things: the width of $\phi_p(z)$, σ_z; the surface density of members of the photometric sample on the sky, Σ_p; and the redshift distribution of the spectroscopic sample, dN_s/dz.

We here ignore the weak cross-correlations induced

by gravitational lensing. These correlations can be predicted directly from the observed galaxy number counts (Scranton et al. 2005). Alternately, it should be possible to iteratively remove the lensing-induced signal once we have an estimate of $\phi_p(z)$ as that will allow us to predict how much cross-correlation with members of s at a given z should result from lensing.

The results of these simulations are shown in Figures 3, 4, and 5. For samples with constant dN_s/dz, if sample variance is negligible, the Monte Carlo simulations find that the errors in determining either $\langle z \rangle$ or σ_z are identical, and can be fit extremely well (to within 1%) by:

$$\sigma = 9.1 \times 10^{-4} \left(\frac{\sigma_z}{0.1} \right)^{1.5} \left(\frac{\Sigma_p}{10} \right)^{-1/2} \left(\frac{dN_s/dz}{25{,}000} \right)^{-1/2}$$
$$\times \left(\frac{4\,h^{-1}\,\mathrm{Mpc}}{r_{0,sp}} \right)^{\gamma} \left(\frac{10\,h^{-1}\,\mathrm{Mpc}}{r_{max}} \right)^{2-\gamma} , \quad (12)$$

where Σ_p is expressed in galaxies per square arcminute. Typical values of γ for both local and $z \sim 1$ galaxy samples are 1.7–1.8 (Zehavi et al. 2005; Coil et al. 2006c).

The scaling of uncertainties with σ_z may be understood as the combination of two effects. First, if the x coordinate of a distribution is rescaled by some factor, errors in quantities proportional to x should be rescaled by the same factor, so it is not surprising that $\sigma \propto (\sigma_z/0.1)$ at least. However, there is an additional factor: when σ_z is smaller, $\phi_p(z)$ is more concentrated about the mean value, so fractional errors in ϕ_p from Poisson statistics are smaller about the peak, leading to the additional factor of $(\sigma_z/0.1)^{0.5}$.

If sample variance is corrected for using the observed fluctuations in dN_s/dz, the uncertainty in determining $\langle z \rangle$ is fit fairly well (to within 20%) by:

$$\sigma = 1.4 \times 10^{-3} \left(\frac{\sigma_z}{0.1} \right) \left(\frac{\Sigma_p}{10} \right)^{-0.3} \left(\frac{dN_s/dz}{25{,}000} \right)^{-1/2}$$
$$\times \left(\frac{4\,h^{-1}\,\mathrm{Mpc}}{r_{0,sp}} \right)^{\gamma} \left(\frac{10\,h^{-1}\,\mathrm{Mpc}}{r_{max}} \right)^{2-\gamma} , \quad (13)$$

while the uncertainty in σ_z proves to be 10% smaller. The dependence of errors upon σ_z and Σ_p is significantly weaker in this scenario. When σ_z is smaller, the true redshift distribution covers a smaller range in z, making the impact of sample variance larger; while when Σ_p increases, Poisson errors decrease but sample variance does not, reducing its effects.

For comparison, the estimated requirements for ambitious future surveys such as SNAP or LSST are $\sigma(\langle z \rangle) \lesssim$ 2-4$\times 10^{-3}$ at $z \sim 1$ (see §1); throughout the remainder of this paper, we will take 3×10^{-3} as a reasonable target for these projects, and indicate it by a dashed line in the relevant figures. Even with one-tenth the surface density assumed in our standard scenario, estimated errors are within this limit. Cross-correlation techniques can meet the calibration requirements of next-generation dark energy surveys.

3.1.1. *Combining Survey Samples*

We now consider more realistic scenarios, where the spectroscopic sample is a combination of real or planned redshift surveys. In general, different redshift surveys
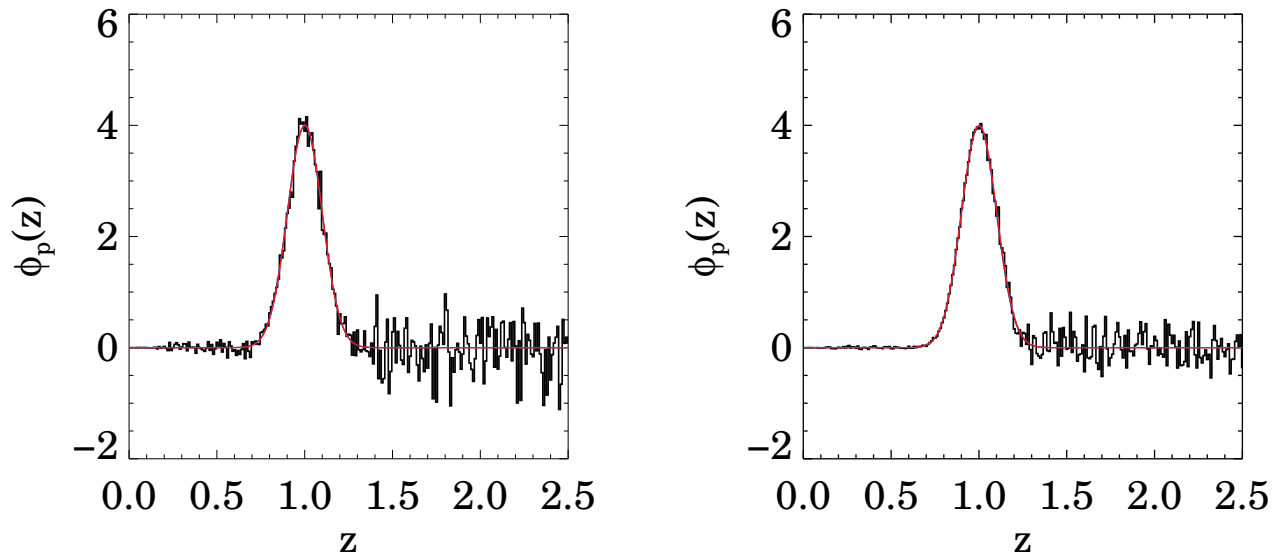
FIG. 2.— Examples of individual Monte Carlo realizations for the recovery of $\phi_p(z)$ using the combinations of current spectroscopic datasets (left) or of current and future datasets (right) shown in Fig. 1. Each realization was generated by randomly drawing from realistic error distributions for the recovery of $\phi_p(z)$ in bins of width $\Delta z = 0.01$. Plotted in blue is the true, input redshift distribution, given by Equation 11 with $\sigma_z = 0.1$. The black histogram shows one realization for the distribution measured using cross-correlation techniques, with realistic errors determined as described in §2.2. Shown in red is the distribution determined from a least-squares fit to the simulated data shown by the black histogram. The recovery is good enough in each case that the blue curve is essentially invisible.

TABLE 1
ASSUMED REDSHIFT SURVEY SAMPLES

Survey name	# of high-confidence z's	z_0[a] or redshift range	Reference
Current Samples			
Sloan Digital Sky Survey (SDSS)	800,000	0.017	Strauss et al. 2002
AGN & Galaxy Evolution Survey (AGES)	10,000	0.09	Kochanek et al. 2004
DEEP2 Galaxy Redshift Survey, EGS	8,200	0.225	Davis et al. 2006
DEEP2 Galaxy Redshift Survey, non-EGS	17,000[b]	0.225[b]	Faber et al. 2007
VIMOS/VLT Deep Survey	10,000[c]	0.27	Le Fèvre et al. 2005
Lyman/Balmer break samples, $1.5 < z < 4$	2,500	$1.5 < z < 4$	Steidel et al. 1999, 2003, 2004
Near-future Samples			
WiggleZ	350,000	$0.25 < z < 1$[d]	Glazebrook et al. 2007, in prep.
PRIMUS	300,000	0.23[e]	Eisenstein et al. 2007, in prep.
zCOSMOS, $I < 22.5$	5,000[c]	0.23	Lilly et al. 2006
zCOSMOS, high-z	2,500[c]	$1.5 < z < 2.5$	Lilly et al. 2006
Lyman/Balmer-break samples, $1.5 < z < 4$	5,000	$1.5 < z < 4$	N/A

[a] Except where redshift ranges are specified, we assume redshift distributions are proportional to $z^2 e^{-z/z_0}$, and so have median redshift given by $2.67z_0$ and mean redshift $3z_0$. We then use published median redshifts or photometric depths to estimate z_0, as described in the text. [b] Outside of the Extended Groth Strip (EGS), DEEP2 uses a color cut complete for $z > 0.75$ ($\sim 50\%$ complete at $z = 0.7$). [OII] 3727 and the 4000-Åbreak leave the DEIMOS spectral window at $z \gtrsim 1.4$, so redshift success is minimal beyond that point. We therefore include only redshift quality=4 objects with $0.7 < z < 1.4$ in this count and the model redshift distribution. [c] For VVDS and zCOSMOS, we take a 20% rate of flag=4 redshifts, following Le Fèvre et al. (2005), and take a full VVDS sample size of 50000 objects. We optimistically assume that the median redshifts given in Le Fèvre et al. (2005) apply for the flag=4 galaxies, though very few of them are at $z > 1$ (Ilbert et al. 2005). [d] The redshift distribution from test WiggleZ observations may be roughly approximated by a Gaussian of mean 0.55 and σ 0.25, truncated at $z = 0.25$ and $z = 1$. [e] We presume that, in addition to being I-band limited, PRIMUS will generally not measure redshifts for objects with $z > 0.95$ (A. Coil, priv. comm.).
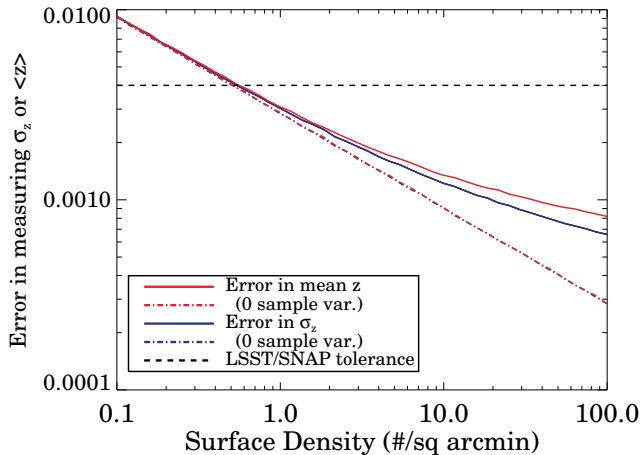
FIG. 3.— Errors in the recovery of the mean redshift, $\langle z \rangle$ (red) or the RMS dispersion of redshifts, σ_z (blue) for objects in a photometric sample versus their surface density in galaxies per square arcminute, Σ_p, as measured in our Monte Carlo simulations. Note that for a single photometric redshift bin drawn from a larger sample, Σ_p is the surface density only for objects in that bin, not for the overall sample. The black, dashed line indicates the estimated maximum error in $\langle z \rangle$ allowable for proposed dark energy surveys using the SNAP satellite or LSST. We assume a spectroscopic sample with $dN_s/dz = 25,000$ (roughly corresponding to current samples at $z \sim 1$) and a true $\phi_p(z)$ having $\sigma_z = 0.1$. If sample variance is negligible, both errors scale as $\sigma \propto \Sigma_p^{-1/2}$; if it has maximal impact, their scaling is weaker, $\sigma \propto \Sigma_p^{-0.3}$. If the photometric sample has very low surface density, larger numbers of redshifts or a narrower redshift distribution than assumed in our standard scenario may be required to meet the requirements of future dark energy surveys.

will be optimized for different depths or redshift ranges; hence, to cover the full possible redshift range of the photometric objects, a combination of different redshift surveys is likely to be used. This does not present any fundamental problems; ξ_{sp} and ξ_{ss} may be determined and $\phi_p(z)$ estimated separately for each dataset; the resulting $\phi_p(z)$ from each sample may be combined with weighted means.

If random errors on each sample's clustering measurements are small ($\lesssim 1\%$), as is generally true in large modern surveys, the net errors on $\phi_p(z)$ should be the same when we simply use the aggregate dN_s/dz as if we measure separately and combine (we test the impact of clustering measurement uncertainties in §4.2). We hence estimate the combined dN_s/dz from current samples and from surveys that have recently begun observations. The samples considered, their estimated median redshifts, and each sample's total number of galaxies are given in Table 1. Where other information is not available, we have used announced or published magnitude limits and the fitting formulae from Coil et al. (2004b) to estimate median redshifts. We assume that except for hard redshift limits set by color cuts or lack of features in spectral windows, all $z < 2$ samples have redshift distributions of the form $z^2 e^{-z/z_0}$, which fits current datasets well (Coil et al. 2004b); for distributions of this form, $z_0 = \mathrm{median}(z)/2.67$. For $z \geq 2$, we assume that dN_s/dz will be flat, the rough consequence of applying a wide variety of high-redshift selections targeted at different z ranges. The estimated combined redshift distributions for current and near-future samples used here are shown
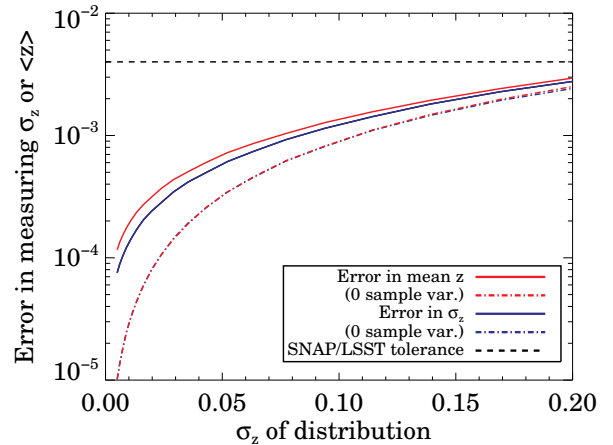

FIG. 4.— Errors in the recovery of $\langle z \rangle$ (red) or σ_z (blue) versus the true value of σ_z, from our Monte Carlo tests. The black, dashed line indicates the estimated maximum error in $\langle z \rangle$ allowable for proposed dark energy surveys using LSST or the SNAP satellite. We assume here a spectroscopic sample with $dN_s/dz = 25,000$ (roughly corresponding to current samples at $z \sim 1$) and a photometric sample with a surface density of 10 galaxies per square arcminute. If sample variance is negligible, both errors scale as $\sigma_z^{3/2}$; if it has maximal impact, their σ_z–dependence is weaker, $\sigma \propto \sigma_z$. In all plotted cases, the errors in measuring the parameters of the redshift distribution are much smaller than required for future dark energy surveys.

in Figure 1.

For the sample of current surveys (all but zCOSMOS, WiggleZ, and PRIMUS from Table 1) we then find:

$$\sigma = 1.2 \times 10^{-3} \left(\frac{\sigma_z}{0.1} \right) \left(\frac{\Sigma_p}{10} \right)^{-0.3}, \qquad (14)$$

while for a reasonable near-future scenario (including only projects that have begun observations), we find:

$$\sigma = 4.1 \times 10^{-4} \left(\frac{\sigma_z}{0.1} \right)^{1.5} \left(\frac{\Sigma_p}{10} \right)^{-1/2}, \qquad (15)$$

corresponding well to the tolerances for future dark energy experiments ($\sigma(\langle z \rangle) \lesssim 3 \times 10^{-3}$, as described in §1). We have assumed that sample variance is negligible for the near-future scenario, as the largest-area projects (SDSS and WiggleZ) plan to survey at least 1000 square degrees each (see 3.4).

3.2. Non-parametric reconstruction

Although these methods are highly successful at reconstructing the parameters of $\phi_p(z)$ given the correct general model, we can also test how well we may recover $\langle z \rangle$ when making no assumptions about $\phi_p(z)$ at all. We hence calculate the recovered mean redshift for each Monte Carlo realization, $\langle z \rangle = \sum z_i \phi_i / \sum \phi_i$, where z_i is the redshift of the ith bin, ϕ_i is the recovered $\phi_p(z)$ in that bin, and we use $\sum$ to indicate summation over all bins i. We then find that, for our standard scenario and averaging over the redshift range $0 < z < 2$, the standard deviation of these mean redshift estimates is:
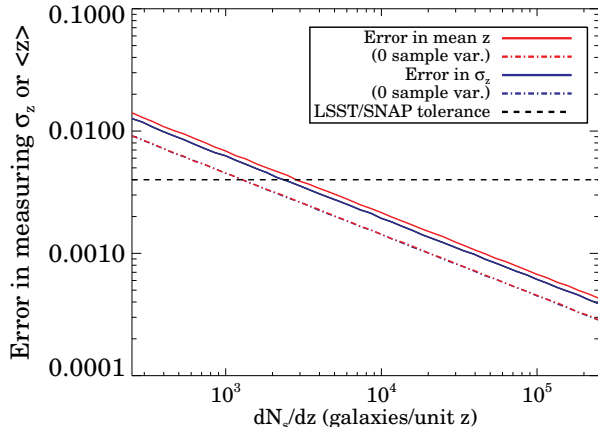
FIG. 5.— Errors in the recovery of $\langle z \rangle$ (red) and σ_z (blue) versus the number of spectroscopic galaxies per unit redshift, dN_s/dz, from our Monte Carlo tests. The black, dashed line indicates the estimated maximum error allowable for proposed dark energy surveys using the SNAP satellite or LSST. We assume here that the photometric sample has a surface density of 10 galaxies per square arcminute and a true $\phi_p(z)$ having $\sigma_z = 0.1$. Regardless of assumptions about sample variance, all uncertainties scale as $(dN_s/dz)^{-1/2}$. If dN_s/dz is small, as at $z > 2$ currently, meeting the tolerances of future dark energy surveys may be problematic. However, accuracy requirements at $z \sim 2$ are in general less restrictive than the $z \sim 1$ tolerance plotted here, as angular diameter distance and lookback time evolve more slowly with redshift at higher z.

$$\sigma(\langle z \rangle) = 6.9 \times 10^{-3} \left(\frac{\Sigma_p}{10} \right)^{-1/2} \left(\frac{dN_s/dz}{25,000} \right)^{-1/2}$$
$$\times \left(\frac{4 \, h^{-1} \, \mathrm{Mpc}}{r_0} \right)^{\gamma} \left(\frac{10 \, h^{-1} \, \mathrm{Mpc}}{r_{max}} \right)^{2-\gamma} \quad (16)$$

if sample variance is corrected using the observed dN_s/dz; the prefactor is 6.4×10^{-3} if sample variance is negligible. These errors would be reduced if the redshift range considered is more limited. We plan to further explore the effectiveness of nonparametric reconstruction of redshift distributions in future work.

3.3. Impact of redshift outliers

The objects that fall within a photometric redshift bin generally are not a pure population. Non-Gaussian photometric errors, e.g. due to contamination by light from overlapping objects, may cause some galaxies to incorrectly be placed in a given bin (our sample p), while in other cases the observed colors of galaxies at very different redshifts may be degenerate, causing $\phi_p(z)$ to be multimodal.

If $\phi_p(z)$ consists of a combination of multiple Gaussians that overlap only minimally, the scalings from Equation 12 should hold for each peak, save that we must replace Σ_p, the total surface density of photometric objects, by $f\Sigma_p$, where f is the fraction of sample p that is associated with a given peak. As the peaks begin to overlap, however, this prescription will fail. We have therefore adapted our Monte Carlo simulations to test the recovery of a distribution function consisting of two Gaussian peaks of equal width and amplitude (centered at redshifts z_1 and z_2) as z_2 approaches z_1; i.e., we employ
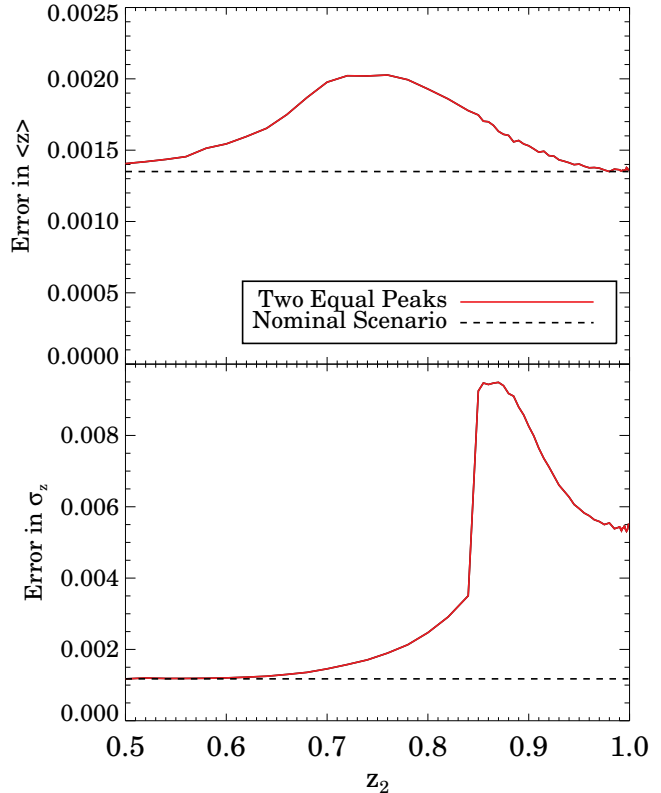

FIG. 6.— Errors in the recovery of the mean redshift of the photometric sample, $\langle z \rangle$ (top), or the Gaussian σ of the redshift distribution, σ_z (bottom), for scenarios where the redshift distribution of the photometric sample $\phi_p(z)$ consists of two Gaussian peaks of equal height, one centered at redshift 1 and the other at redshift z_2. We assume for these simulations that the photometric sample has a surface density of 10 galaxies per square arcminute, that each peak has Gaussian $\sigma_z = 0.1$, and that the spectroscopic sample has $dN_s/dz = 25,000$. Although recovery of σ_z is degraded when the two peaks are not resolved, measurement of $\langle z \rangle$ improves compared to the intermediate regime. Results are qualitatively similar if the true σ_z is changed.

a distribution function $\phi_p(z) = 1/(8\pi\sigma_z^2)^{-1/2} (\exp(-(z - z_1)^2/2\sigma_z^2) + \exp(-(z - z_2)^2/2\sigma_z^2))$. For convenience, we take $z_1 = 1$ and $z_2 < z_1$, though behavior should not depend strongly on these choices. The uncertainties in measuring $\langle z \rangle$ and σ_z for $\sigma_z = 0.1$ are shown in Fig. 6; we obtain qualitatively similar results for $\sigma_z = 0.05$ or 0.2.

Matching the predictions above, when the two peaks overlap minimally, the error in the recovered mean, $\langle z \rangle = (\langle z_1 \rangle + \langle z_2 \rangle)/2$, is equal to the sum in quadrature of the errors in determining each peak's position, divided by two. This is equal to the error obtained for a single Gaussian peak (since $f = 0.5$, so the errors in $\langle z_1 \rangle$ and $\langle z_2 \rangle$ are $\sqrt{2}$ times larger than the single-peak error, but the error of the mean of the two quantities is $1/\sqrt{2}$ as large as the error in one) As z_2 approaches 1, errors reach a maximum of $\sim 1.5\times$ the minimum value when $(z_2 - z_1) \approx 2.5\sigma_z$, and then decrease monotonically, approaching the minimum value again when $(z_2 - z_1) << \sigma_z$. The behavior of the uncertainty in recovering σ_z is more complex, rising rapidly (by $> 5\times$) when the two peaks are unresolved $((z_2 - z_1) \lesssim 1.2\sigma_z)$; fortunately, dark energy experiments are generally less affected by errors in σ_z

than $\langle z \rangle$ (Ma et al. 2006).

Alternatively, we can consider a scenario where redshift outliers share the same mean as the overall sample, but have a broader σ_z. For this, we consider a distribution function $\phi_p(z) = (1 - f_{outlier})(2\pi\sigma_1^2)^{-1/2} \exp(-(z - z_0)^2/2\sigma_1^2) + f_{outlier}(2\pi\sigma_2^2)^{-1/2}\exp(-(z-z_0)^2/2\sigma_2^2)$); that is, the sum of two Gaussians, with total probability of (1-$f_{outlier}$) and $f_{outlier}$ and standard deviations σ_1 and σ_2, respectively, but sharing the same mean, z_0. As before, we produce Monte Carlo realizations for the recovery of such a distribution with cross-correlation techniques, and then fit for $f_{outlier}$, z_0, σ_1, and σ_2. For an initial guess for each realization, we take random values of z_0, σ_1, and σ_2 with an RMS dispersion of 10% about their true value, and a value of $f_{outlier}$ with RMS dispersion 20% of its true value.

For convenience, we simulate distributions $z_0 = 1$ and $f_{outlier} = 0.1$, a realistic value for faint samples (see, e.g., Ilbert et al. 2006). We set $\sigma_1 = 0.05$, 0.1, or 0.2, and investigate the recovery of $\langle z \rangle$, σ_2, and the net σ of the distribution, $((1 - f_{outlier})^2\sigma_1^2 + f_{outlier}^2\sigma_2^2)^{1/2}$, as a function of σ_2. Results for $\sigma_1 = 0.1$ are shown in Fig. 7.

The conclusions are similar for all three values of σ_1. As seen in the top panel of Fig. 7, the presence of a small fraction of objects with a greater σ_z causes a correspondingly small (10-20%) degradation in the recovery of $\langle z \rangle$. This is not be a great surprise. Since the two Gaussians are required to have the same mean, the principal impact on $\langle z \rangle$ is due to the broader effective σ of the distribution; as we found before, reconstruction of $\langle z \rangle$ grows poorer when σ_z is greater.

When σ_2 has a value close to that of σ_1, differentiating the two components becomes difficult; therefore, errors in both σ_1 and σ_2 are greater for low values of σ_2. Errors in $f_{outlier}$, too, are correspondingly higher in this regime. The result is that the net σ of the distribution can be degraded substantially where $\sigma_2 \approx 2\sigma_1$, by more than a factor of 3. For smaller values of σ_2 than this, analysis becomes difficult, as the fitting routine will trade off which Gaussian component corresponds to which piece of the distribution. As σ_2 increases past the point where the two distributions become distinguishable, the error in recovering σ_2 also steadily increases, consistent with the increase in errors in σ_z as σ_z increases for our standard scenario. However, $f_{outlier}$ is better determined as σ_2 goes up, such that the RMS error in the net σ decreases monotonically as σ_2 increases.

3.4. *Impact of sample variance*

In §3.1, we established the minimum and maximum impact sample variance will have on measurements using cross-correlation techniques. Here, we attempt to establish quantitatively in which regimes a correction using the observed dN_s/dz, as described in §2.3, will mitigate the additional errors caused by sample variance, and in which regimes cosmic variance is negligible and such a correction is inadvisable.

To do so, we have performed another set of Monte Carlo simulations, in which we have added in quadrature to the result of Equation 9 an additional uncertainty corresponding to the fluctuations in the count of an unbiased tracer of dark matter in each bin due to sample variance when producing each realization. We assume that
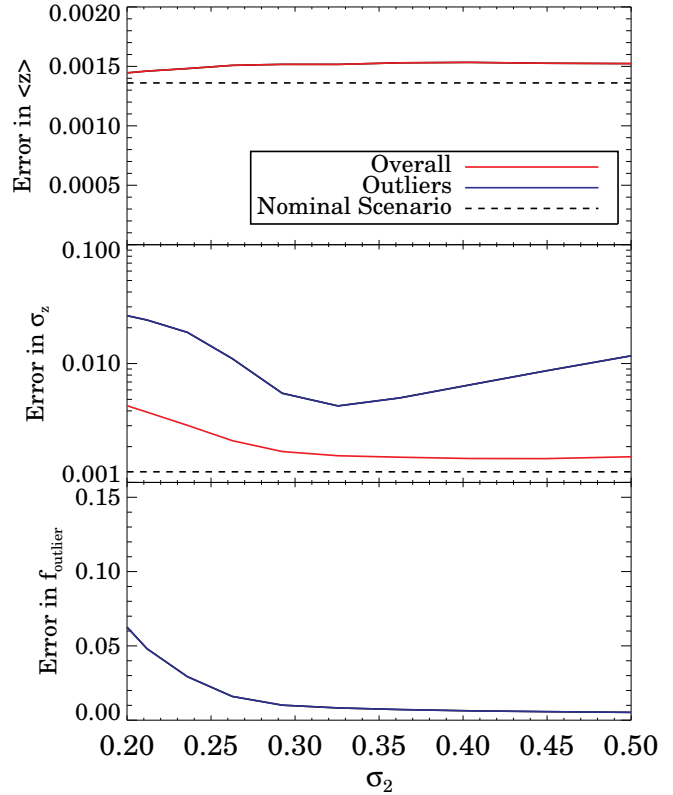

FIG. 7.— Errors in the recovery of the mean redshift of the photometric sample, $\langle z \rangle$ (top), the Gaussian σ of a redshift distribution, σ (middle), or the outlier fraction $f_{outlier}$ (bottom) for scenarios where the redshift distribution of the photometric sample $\phi_p(z)$ consists of two Gaussian peaks with equal mean, one having integral (1-$f_{outlier}$) and $\sigma = 0.1$, and the other having integral $f_{outlier}$ and $\sigma = \sigma_2$. All curves are plotted as a function of the Gaussian σ of the outlier redshift distribution, σ_2. We assume for these simulations that the photometric sample has a surface density of 10 galaxies per square arcminute and that the spectroscopic sample has $dN_s/dz = 25,000$. Red curves indicate the error in the overall mean or net σ of the distribution; blue curves indicate the error in recovering the width of the outlier distribution, σ_2 or the outlier fracton. The dashed line in each case shows what the errors would be if $f_{outlier} = 0$. In the top panel, only one curve is shown, as the two Gaussians are required to have equal mean while fitting. Results are qualitatively similar if the value of σ_1 is 0.05 or 0.2.

a total of N_{fields} independent (i.e., widely separated) fields are covered by the spectroscopic samples, with each field having the same dimensions on the sky. For simplicity, we consider only two field geometries here: either $1 \deg \times 0.5 \deg$ (correspondingly roughly to the sizes of the independent fields surveyed by current deep surveys such as DEEP2 and the VIMOS-VLT Deep Survey [VVDS]), or $2 \deg \times 2 \deg$ (corresponding to proposed future surveys). We calculate the uncertainties from sample variance using the methods of Newman & Davis (2002), and find that the expected fractional root-mean-square (RMS) variations in counts of an unbiased tracer are 48% for the smaller field size and 22% for the larger over $\Delta z = 0.01$.[3] Although the fields are 8× larger in area for $2 \deg \times 2 \deg$ fields, the fluctuations in counts due to sample variance are only $\sim 2.2\times$ smaller; the power spectrum remains non-negligible on degree scales

[3] IDL code is available at `http://astro.berkeley.edu/~jnewman/research.html`.

at $z \sim 1$. We use the same parameters as for our standard scenario here ($\Sigma_p = 10$ galaxies/square arcminute, $dN_s/dz = 25,000$, and $\sigma_z = 0.1$).

For these calculations, we treat the fluctuations in counts from sample variance in successive redshift bins as independent; this assumption is fairly good (e.g. for $1 \deg \times 0.5 \deg$ field sizes and $\Delta z = 0.01$, the covariance between counts in adjoining redshift bins is roughly 16% of the total variance; for $\Delta z = 0.05$, it is only 3%). The root-mean-square errors in recovering $\langle z \rangle$ as a function of N_{fields} are shown in Figure 8. For $1 \deg \times 0.5 \deg$ fields, the uncertainty when sample variance fluctuations are not corrected for is worse than the errors resulting from using the observed dN_s/dz to make corrections so long as $N_{fields} \lesssim 15$. For the larger field size, sample variance is worse than the uncertainties in the correction only for $N_{fields} \lesssim 3$. If more fields than this are surveyed, there is no advantage to using fluctuations in the spectroscopic redshift distribution to correct for sample variance. A number of fields roughly $4 - 5\times$ greater is required for sample variance to have completely negligible impact.

3.5. Covariance of Sample Variance

Our Monte Carlo simulations assumed that fluctuations in counts due to sample variance are independent between all redshift bins. However, this is of course not the case; e.g., high peaks will tend to cluster together. We therefore must assess to what degree this covariance will worsen the reconstruction of $\phi_p(z)$.

A simple way to test this is to determine how errors in the parameters of $\phi_p(z)$ change when the redshift bin size used is altered. The Monte Carlo simulations described above incorporate the total sample variance within whatever bin size is used; e.g., the errors from sample variance used if bin sizes double are not simply $1/\sqrt{2}$ as large as before. One caveat for this test is that recovery of the parameters of $\phi_p(z)$ may be affected by discreteness effects (the mean of $\phi_p(z)$ within a bin is not identical to the value of $\phi_p(z)$ at the bin's center, although it is treated as such in fitting) when redshift bins become large, independent of any sample variance effects.

For our standard scenario ($\sigma_z = 0.1$, 4 independent fields, $\Sigma_p = 10$), we find that errors in $\langle z \rangle$ and σ_z rise steadily as bin size increases if the observed dN_s/dz is not used to correct for sample variance fluctuations, reaching 25% larger values for $\Delta z = 0.1$ bins than for $\Delta z = 0.01$. If dN_s/dz is used for corrections, however, errors in reconstruction are flat with bin size to 1% or better; this is no surprise, as in such scenarios, we are limited by Poisson uncertainties in the correction (which should have no covariance between bins) rather than the sample variance itself.

Because of the possibility of discreteness effects, we have investigated the effects of the covariance of sample variance with a model that may be employed even for small bin sizes. We proceed by showing that the dominant effect is a covariance only between successive redshift bins, with larger-scale effects being comparatively negligible; and then show that incorporating this leading-order effect causes only a modest degradation in the recovery of $\langle z \rangle$ and σ_z.

Specifically, let us suppose that the fractional fluctuation in a count from sample variance in the ith redshift bin, which we will label s_i, is covariant only with the

fluctuations in the neighboring bins (s_{i-1} and s_{i+1}. We also assume that the RMS variations from sample variance in each bin are equal – this holds to high accuracy to high accuracy for bins of constant Δz (based upon tests with the QUICKCV code from Newman & Davis 2002) – and that the covariance between bins similarly does not depend on z; we only need these assumptions to hold locally (i.e., for small Δ_i). Even if present, small asymmetries in the impact of sample variance with redshift would affect this test only modestly, however, as they would simply mean that the RMS impact is slightly less on one side than another, but the overall effect of the covariance would remain largely unchanged. If the impact of covariance proved to be large, this might begin to make a quantitative difference, and it might be necessary to include such effects.

As an additional caveat, because it calculates in real space, QUICKCV does not account for the fact that correlation functions and the power spectrum are asymmetric in redshift space. Peculiar velocities will cause the true covariance between successive bins to be overestimated by the procedure below, as their dominant effect on large scales is the "Kaiser infall" (Kaiser 1987), which causes large structures to appear collapsed along the line-line-of-sight in redshift space. This is particularly the case for optically-selected samples at high redshift, which are biased towards intrinsically blue galaxies and have only very weak "Fingers of G-d" (Coil et al. 2007); or if the most nonlinear, sub-Mpc scales are excluded from cross-correlation analyses, as suggested in §2.2. The net effect of redshift-space distortions is that sample variance fluctuations will be more confined to a single redshift bin than one would expect from a real-space calculation, so our estimated errors from this model will be conservative (i.e., biased high).

Given our assumptions, we may presume that for each bin there is an underlying, 'hidden' variable, s_i', which has Gaussian random variations completely independent of the adjoining bins; and we can write

$$s_i = (1 - 2w)s_i' + ws_{i-1}' + ws_{i+1}', \qquad (17)$$

where w is some unknown weight factor. We take the RMS variation for each of the uncorrelated s_i' to be given by the variable σ_u (for uncorrelated), which by assumption is the same for all i.

Given this model and standard propagation of errors, it is possible to predict the net fluctuation due to sample variance for a single bin of width Δz, $2\Delta z$, $3\Delta z$, etc. The first three of these are:

$$\sigma_{\Delta z}^2 = (6w^2 - 4w + 1)\sigma_u^2 \qquad (18)$$

$$\sigma_{2\Delta z}^2 = \frac{1}{4}(4w^2 - 4w + 2)\sigma_u^2 \qquad (19)$$

$$\sigma_{3\Delta z}^2 = \frac{1}{9}(4w^2 - 4w + 3)\sigma_u^2. \qquad (20)$$

Therefore, given numerical predictions for $\sigma_{\Delta z}$ and $\sigma_{2\Delta z}$ given the full power spectrum (predictions which we take from QUICKCV), it is possible to solve for w and σ_u in this model. We can then use the same code to determine $\sigma_{3\Delta z}$, and compare it to the prediction of the model, in order to assess the model's effectiveness at incorporating the impact of the covariance of sample variance.

We find that this simple model is highly effective. For bin sizes (i.e. Δz) ranging from 0.003 to 0.1, ignoring

the covariance of sample variance completely (so that $\sigma_{3\Delta z} = \sigma_{\Delta z}/\sqrt{3}$) underpredicts the RMS fractional variation in a count in a bin of width $3\Delta z$ by anywhere from 2% (for $\Delta z = 0.1$) to 23% (for $\Delta z = 0.003$); for our standard $\Delta z = 0.01$ bins, the underprediction is 15% (all tests are for a 1 deg × 1 deg field with central redshift $z = 1$). If we employ the model described above, however, the underprediction ranges from below 0.1% (for $\Delta z \geq 0.025$) to 3% (for $\Delta z = 0.003$); for $\Delta z = 0.01$ the prediction is off by 0.6%. Clearly, the vast majority of the effect of the covariance of sample variance is described simply by a covariance between successive redshift bins. This holds true even if we had tested over a larger z range; e.g., for $\Delta z = 0.01$, the model correctly predicts the net variance over $5\Delta z$ to within 1.5%, or over $10\Delta z$ to within 2.1%. However, in the latter case, the assumption that all the bins are statistically independent performs even better, matching to 0.4%; so this model will actually overpredict the impact of the covariance of sample variance on large scales.

The form of equation 17 is particularly convenient for incorporation into our Monte Carlo tests. Instead of simply randomly drawing the fractional fluctuation from sample variance in a given bin, s_i, as before, instead we can draw the uncorrelated s_i', and then construct the correlated s_i from the uncorrelated s_i' in making each Monte Carlo realization; we need only predict w and σ_u (which we do using QUICKCV). The impact of adding this covariance to our models may be seen in Fig. 8; errors are increased by roughly 30% in the worst case (for $N_{fields} = 1$), but by < 15% for $N_{fields} > 10$, the regime in which correcting for sample variance fluctuations with the observed dN_s/dz becomes ineffective. For our scaling scenario ($N_{fields} = 4$), errors are degraded by 27%, slightly worse than the difference between $\Delta z = 0.01$ and 0.10. We expect that any corrections from the much smaller, larger-range covariance would be considerably less than this; hence, we conclude that the covariance of sample variance has relatively minor impact on our results. The prefactor of 9.1×10^{-4} in 12 becomes 1.2×10^{-3} when this covariance is accounted for, still well within SNAP and LSST requirements.

The impact is even smaller for our standard scenario, for which the observed dN_s/dz is used to correct for sample variance in each redshift bin. In that case, the Poisson uncertainty in this correction is far greater than the covariance from sample variance, and the latter becomes totally negligible; hence, all the major results of this paper are unaffected by this covariance.

4. POSSIBLE SYSTEMATICS

We now consider the impact of a variety of effects that violate the simple assumptions underlying our basic scenario. We will treat these errors analytically wherever possible. We summarize the results of this section in Table 2.

4.1. Evolution in bias

Inspection of Equation 4 shows that to transform the cross-correlation signal w_{sp} into $\phi_p(z)$, we require knowledge of the parameters of ξ_{sp}, namely γ and $r_{0,sp}$, as a function of redshift. As described in §2.1, however, w_{pp} provides constraints only on the mean clustering of the
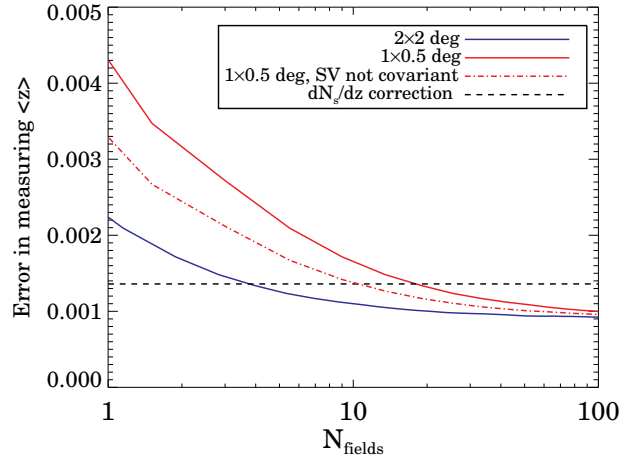

FIG. 8.— Errors in the recovery of $\langle z \rangle$ versus the number of independent fields surveyed for two different field geometries (red and blue curves), resulting from Monte Carlo tests in which sample variance errors are added to our standard scenario ($\Sigma_p = 10$ galaxies per square arcminute, $\sigma_z = 0.1$, $dN_s/dz = 25,000$). As the number of fields increases or if larger fields are used, errors from sample variance decrease. However, an upper limit on the practical impact of sample variance is set by the black, dashed curve, which indicates the errors if the observed dN_s/dz distribution is used to correct for the density fluctuations in each redshift bin. The errors after this correction are set by Poisson statistics from the spectroscopic sample, rather than by clustering. For current redshift samples, applying such a correction (as assumed in the preceding plots) is favored; however, WiggleZ and subsequent surveys should cover ∼ 1000 square degrees each, so sample variance will affect them only minimally. The red, dot-dashed curve shows what the errors would be if sample variance were not covariant between redshift bins; see §3.5.

photometric sample. What is the impact on the derived ϕ_p if these parameters evolve?

We test this by assuming that the net change in ξ_{sp} with redshift is due only to changes in the linear bias of of the photometric sample, b_p. We further assume that the evolution of the bias is linear; i.e., $b_p(z) = b_p(1) + (db/dz)(z - 1)$ with constant db/dz, and that db/dz is small compared to $b_p(1)$. As usual, we take $\phi_p(z)$ to be given by Equation 11, and adopt the simple-biasing assumption $\xi_{sp} = (\xi_{ss}\xi_{pp})^{1/2}$, so $\xi_{sp} \propto (b_p(z)/b_p(1))^{1/2}$. Then, if b_p varies but w_{sp} is interpreted with a constant b_p, to leading order in db/dz the measured $\langle z \rangle$ will be off by:

$$\Delta(\langle z \rangle) = \frac{\int z \left(\frac{b_p(z)}{b_p(1)}\right)^{1/2} g(z)dz}{\int \left(\frac{b_p(z)}{b_p(1)}\right)^{1/2} g(z)dz} - \frac{\int z\, g(z)\, dz}{\int g(z)\, dz}$$

$$\approx \frac{1}{\sqrt{2\pi}\sigma_z} \int -z \frac{db/dz}{2\, b_p(1)}(z - 1)\; e^{-\frac{(z-1)^2}{2\sigma_z^2}}\, dz$$

$$= \frac{db/dz}{2\, b_p(1)}\sigma_z^2, \tag{21}$$

applying the linear approximation $(1+\epsilon)^{1/2} \approx (1+\epsilon/2)$.

By comparing the observed galaxy clustering to the predicted clustering of dark matter in a $\sigma_8 = 0.9$ model from Smith et al. (2003), Coil et al. (2006c) estimate that the linear bias of ∼ L_* galaxies at $z \sim 1$ is 1.48±0.04. Applying the same methods at $z = 0$, the correlation measurements of Zehavi et al. (2005) correspond to a linear bias for L_* galaxies in the Sloan Digital Sky

Survey of 1.0, differing by $\lesssim 0.5 \pm 0.1$. We therefore take 0.5 to be a reasonable upper limit on $[(db/dz)/b_p(1)]$. If the bias of the photometric sample evolves but we do not take that into account in our analysis, we would make a systematic error of at most 0.0025 for $\sigma_z = 0.1$, pushing the tolerances of future dark energy projects. In actuality, however, we do possess information on bias evolution (e.g. by measuring the mean b_p over successive photometric bins), so it is very unlikely this effect would lead to an error that large.

It is worth noting that spectroscopic surveys with multiple selection techniques (e.g. by a variety of different color cuts) may select objects with very different bias according to the technique used. This could, for instance, lead to jumps in bias at the redshifts where a given subsample becomes relevant or irrelevant. In that case, it would be most effective to split the survey into its constituent subsamples, treating the clustering/bias of each subsample separately.

4.2. Errors in measuring autocorrelations of spectroscopic samples

A second potential issue is systematic errors in measuring the autocorrelation function of the spectroscopic sample, ξ_{ss}. Although random errors in modern surveys are small, it is very difficult to reduce systematic errors below 1-2%. These systematic errors will generally cause ξ_{ss} to be over- or under-estimated similarly at all z covered by a given survey. To assess the impact of these systematics, we test their worst-case impact using our standard $\phi_p(z)$ distribution.

Thus, we assume that there are only two redshift surveys; that one survey covers the complete $z \leq 1$ regime, and another $z > 1$; and that the ξ_{ss} measurement for each survey may suffer an unknown systematic of RMS fractional amplitude σ_{sys}; i.e., if $\sigma_{sys} = 0.02$, we expect the measured amplitude of ξ_{ss} (for similar galaxies at the same redshift) to vary by 2% from survey to survey. Thus, one half of the reconstructed $\phi_p(z)$ will have an amplitude differing from the other by a factor r_{sys} drawn from a distribution with RMS σ_{sys} and mean 1 (the difference is σ_{sys} rather than $2\sigma_{sys}$ due to the simple-biasing assumption and linear approximation to the square root, as in §4.1). Then:

$$\Delta(\langle z \rangle) = \frac{\left[\int_{-\infty}^{1} z \, g(z) \, dz + r_{sys} \int_{1}^{\infty} z \, g(z) \, dz \right]}{\left[\int_{-\infty}^{1} g(z) \, dz + r_{sys} \int_{1}^{\infty} g(z) \, dz \right]} - 1$$

$$= \frac{r_{sys} - 1}{r_{sys} + 1} \frac{\sqrt{2}\sigma_z}{\sqrt{\pi}}$$

$$\approx \frac{r_{sys} - 1}{2} \frac{\sqrt{2}\sigma_z}{\sqrt{\pi}}, \qquad (22)$$

taking σ_{sys} to be small, so $r_{sys} \approx 1$. Propagating errors, we then find that in a worst-case scenario, $\langle z \rangle$ will have a root-mean-square bias of $(1/2\pi)^{1/2}\sigma_{sys}\sigma_z$, or more conveniently, $8.0 \times 10^{-4}(\sigma_{sys}/0.02)(\sigma_z/0.1)$, well within estimated tolerances for SNAP and LSST.

4.3. Field-to-field zero point variations

A third factor not considered in our standard scenario is spatial variation in the effective zero points of the photometry used to define the photometric sample, p (due to seeing, calibration issues, etc.). Note that random photometric errors or absolute zero point uncertainties have no effect on redshift distributions measured from cross-correlations; the method will empirically determine $\phi_p(z)$ for whatever falls in a given photometric redshift bin, regardless of whether a given object is put in that bin due to errors or because it rightfully belongs there. Instead, the principal impact of zero point errors will be changes in the effective depth of the sample between separately calibrated patches; an error of Δm magnitudes will lead to a fractional error in number counts of $N^{-1}dN/dM \Delta m$. For R-band-limited samples, $d(\log_{10} N)/dM = N^{-1}/(\ln 10)dN/dM \approx 0.36$ (Coil et al. 2004b), so the fractional variation in the surface density of objects in the photometric sample (Σ_p), will be approximately $0.83\,\sigma_{zp}$, where σ_{zp} is the RMS variation of the photometric zero point in magnitudes. The logarithmic slope of galaxy number counts is larger in B (~ 0.5) and slightly smaller for I (~ 0.33), leading to modestly different prefactors for these cases.

Zero point variations will impact the errors in determining $\phi_p(z)$ in two ways. The first is that, if different spectroscopic surveys cover regions with different photometric zero points, the cross-correlation signal will be artificially boosted or decreased for each survey as σ_{zp} varies, since the overall value of Σ_p used for normalization will be not quite appropriate for the effective magnitude limit in each patch of sky. Again, we consider a worst-case scenario, where one survey covering one set of N_{patch} independent calibration patches is used to reconstruct $\phi_p(z)$ at $z \leq 1$, and another covering a separate set of N_{patch} patches is used for $z > 1$.

We may again apply the results of Equation 22. Since the fractional error in w_{sp} for a single survey will be $0.83\,\sigma_{zp}\,N_{patch}^{-1/2}$ (as we are averaging over N_{patch} calibration patches), the RMS variation in r_{sys} (the ratio of the reconstructed $\phi_p(z)$ at $z > 1$ to $z < 1$) should be $\sqrt{2} \times 0.83\,\sigma_{zp}\,N_{patch}^{-1/2}$. Thus, the worst-case RMS error in a measurement of $\langle z \rangle$ due to zero point variations is $2.3 \times 10^{-3}\,(\sigma_{zp}/0.01)\,(N_{patch}/4)^{-1/2}\,(\sigma_z/0.1)$, within SNAP and LSST tolerances. Specifications for zero point variations are generally smaller than this (e.g. 0.005 mag RMS zero point variation for LSST; cf. Burke et al. 2006), and that ongoing redshift surveys should cover many independently-calibrated patches of sky, so zero point variations are likely to have even smaller impact.

The second effect of zero point variations will be to increase the fluctuations in counts beyond those expected for Poisson errors (as assumed in Equation 9), even if there is only one redshift survey. However, for reasonable scenarios, this is minor; even in a very conservative scenario, with $dN_s/dz = 25,000$, $\Sigma_p = 10$ galaxies per square arcmin, only 3 independently calibrated patches of sky surveyed, and RMS zero point errors of 0.05 mag, errors in mean z and σ_z from Monte Carlo tests increase by a fraction of a percent of their value when this effect is added to our standard scenario.

4.4. Errors in assumed cosmology

As seen in Equation 5, transforming $\tilde{w}(z)$ to $\phi_p(z)$ requires knowledge of the volume element, $dV/dz \, d\Omega = d_A(z)^2 \, dl/dz$. Because the scale radii used are expressed

in h^{-1} Mpc, all scalings with the Hubble parameter cancel out; but $dV/dz\,d\Omega$ will depend on other cosmological parameters as well. Similarly to the preceding sections, we investigate this by assessing the error in $\langle z \rangle$ that will result from assuming a mistaken cosmology. For convenience, we consider only spatially flat, quintessence + cold dark matter models characterized by a matter density Ω_m and dark energy equation-of-state parameter w; these are 0.3 and -1, respectively, for our standard scenario.

We then determine the error in $\langle z \rangle$ that occurs if the cosmology differs from the standard in one of these parameters, but $\tilde{w}(z)$ is interpreted using the standard cosmology. Let $V_{assumed}(z)$ and $V_{true}(z)$ denote the values of $d_A(z)^2\,dl/dz$ for the assumed cosmological model and the true cosmology, respectively. Then mistaken assumptions will lead to an error in $\langle z \rangle$ given by:

$$\Delta\langle z\rangle = \frac{\int_0^\infty z\,V_{assumed}(z)\,g(z)\,dz}{\int_0^\infty V_{assumed}(z)\,g(z)\,dz} - \frac{\int_0^\infty z\,V_{true}(z)\,g(z)\,dz}{\int_0^\infty V_{true}(z)\,g(z)\,dz}. \tag{23}$$

For $\sigma_z \lesssim 0.3$, the effect of varying Ω_m may then be approximated well by $\Delta\langle z\rangle = 4.2 \times 10^{-4}(\sigma_z/0.1)^2(\Delta\Omega_m/0.03)$, where we normalize to a 10% variation in Ω_m (comparable to errors from WMAP; Spergel et al. 2006). The effects of varying w are much less symmetric about our standard scenario, so we consider $w < -1$ and $w > -1$ separately. The impact of errors in cosmology are stronger in the former case: approximately $\Delta\langle z\rangle = 7\times 10^{-5}(\sigma_z/0.1)^{1.9}(\Delta w/0.1)$, where $\Delta w = w - 1$. For $w > -1$, $\Delta\langle z\rangle$ has a turning point at $w \approx -0.93$; the value of $\Delta\langle z\rangle$ at that turning point is approximately $1.7 \times 10^{-5}(\sigma_z/0.1)^{1.9}$. Thus, it is more important to constrain Ω_m than w when using cross-correlation techniques to determine $\phi_p(z)$, but regardless, the cosmological uncertainties are small compared to the requirements of proposed dark energy surveys.

5. CONCLUSIONS

In this paper, we have described a new method for recovering the redshift distribution of objects in a photometric sample by measuring their angular cross-correlations with objects in redshift survey samples as a function of spectroscopic z. This technique does not require that spectroscopic samples resemble the photometric sample in galaxy properties (such as luminosity) or clustering amplitude. We have demonstrated that in realistic scenarios, the redshift distributions of photometric samples may be determined to the precision required by proposed dark energy experiments with this technique. We conclude here by discussing what can be done to optimize future redshift survey datasets to facilitate applications of cross-correlation techniques.

• First, it is apparent from Figure 1 that there are two redshift regimes that are currently much more poorly sampled than others: $0.2 \lesssim z \lesssim 0.7$, and $z > 1.4$. Concentrating future survey efforts on these regimes would be of great benefit for application of cross-correlation methods. Efforts to cover this lower-redshift gap are already underway.

• We emphasize that, although high redshift precision is not requisite for the spectroscopic sample – e.g., ideal

photometric redshifts with $\sigma_z = 0.01$ would be useable for determining the true redshift distribution for a sample with $\sigma_z \gtrsim 0.1$ – it is vitally important that the **purity** of the redshifts be very high. Otherwise, redshift outliers in the spectroscopic sample will cause distortions in the recovered $\phi_p(z)$. Of course, the same holds true for any photometric redshift calibration technique; if false redshifts are used to calibrate a redshift distribution, it is quite likely to be biased in some way.

In most surveys, high-purity redshifts are only obtained for a fraction of the sample. For instance, in the DEEP2 Galaxy Redshift Survey, roughly 80% of successful redshifts (i.e., $\sim 55\%$ of targeted galaxies) fall within the highest purity class, which has a $\sim 0.5\%$ error rate (based on tests with the > 2000 objects observed multiple times), while the remainder have an error rate of roughly 5% (Faber et al. 2006, in prep.). As another example, in the VIMOS-VLT Deep Survey, $\sim 20\%$ of galaxies targeted yield a redshift in their highest (99%) redshift confidence category; almost all of those high-confidence objects have $z < 1$ (Ilbert et al. 2005). In contrast, in the Sloan Digital Sky Survey, which samples local galaxies with much higher signal-to-noise spectra, almost all galaxies yield an accurate redshift (Schlegel et al., in prep.).

In general, the higher signal-to-noise spectra within a sample will generally yield a greater rate of secure redshifts; i.e., the brightest galaxies will dominate samples of high-confidence redshifts. This can be a problem for direct calibration of photometric redshifts, but has little effect on cross-correlations; the spectroscopic sample need not include high-confidence redshifts of faint galaxies, so long as a sufficient number of brighter galaxies at the same redshift are included. This allows larger, shallower surveys to be used to calibrate redshift distributions even of very faint photometric redshift samples.

• Cross-correlations can be analyzed more simply if nonlinearities have minimal impact; e.g., because the spectroscopic sample has relatively weak biasing and scale-dependence in its bias. This suggests that blue, star-forming galaxies may be more useful for this technique than red, early-type galaxies (Zehavi et al. 2005; Coil et al. 2004a; Coil et al. 2007, in prep.). However, red sequence galaxies can yield relatively high-precision photometric redshifts; they therefore may make up for this disadvantage by sheer abundance. Quasars might make an attractive population to use at high redshift given their high luminosity, but the difficulty of measuring their autocorrelations with precision (Coil et al. 2006b) and the possibility that quasars may affect the evolution of nearby galaxies may make their use to measure cross-correlations over any but the largest scales difficult.

Above all, it is important that future photometric dark energy experiments overlap with spectroscopic surveys on the sky. Without this, cross-correlation measurements are impossible. These measurements are not only useful for determining redshift distributions, but also will allow correlation functions to be measured down to much lower luminosities than can be reached spectroscopically. Furthermore, the availability of more photometry in fields with spectroscopy can improve our understanding of the spectroscopic samples by broadening spectral energy distribution measurements; having pho-

TABLE 2
Summary of Random and Systematic Errors

Error type	Corresponding uncertainty in $\langle z \rangle$
Random errors	$1.5 \times 10^{-3} \, (\sigma_z/0.1)^1 \, (\Sigma_p/10)^{-0.3} \, ((dN_s/dz)/25,000)^{-1/2}$ [a]
Random errors, sample variance negligible	$1.0 \times 10^{-3} \, (\sigma_z/0.1)^{1.5} \, (\Sigma_p/10)^{-1/2} \, ((dN_s/dz)/25,000)^{-1/2}$
Not accounting for evolution in bias	$2.5 \times 10^{-3} [db/dz/b_p(1)]/0.5 \, (\sigma_z/0.1)^2$
Systematic errors in autocorrelation measurements	$< 8.0 \times 10^{-4} \, (\sigma_{sys}/0.02) \, (\sigma_z/0.1)$
Field-to-field zero point variations	$< 2.3 \times 10^{-4} \, (\sigma_{zp}/0.01) \, (N_{patch}/4)^{-1/2} (\sigma_z/0.1)$
Errors in assumed Ω_m	$4.2 \times 10^{-4} (\sigma_z/0.1)^2 \, (\Delta\Omega_m/0.03)$
Errors in assumed w	$< 7 \times 10^{-5} \, (\sigma_z/0.1)^{1.9} \, (\Delta w/0.1)$

[a] Throughout this table we give the surface density of the photometric sample, Σ_p, in galaxies per square arcminute.

tometry in five bands for each galaxy in the SDSS spectroscopic sample has been a great boon to studies of galaxy properties. The synergies between photometry and spectroscopy are great, and determination of redshift distributions from cross-correlations is only one of many applications, though a vital one.

REFERENCES

Abraham, R. G., Glazebrook, K., McCarthy, P. J., Crampton, D., Murowinski, R., Jørgensen, I., Roth, K., Hook, I. M., Savaglio, S., Chen, H.-W., Marzke, R. O., & Carlberg, R. G. 2004, AJ, 127, 2455

Albrecht, A., Bernstein, G., Cahn, R., Freedman, W. L., Hewitt, J., Hu, W., Huth, J., Kamionkowski, M., Kolb, E. W., Knox, L., Mather, J. C., Staggs, S., & Suntzeff, N. B. 2006, ArXiv Astrophysics e-prints

Burke, D., Axelrod, T., Claver, C., Frank, J., Geary, J., Gilmore, K., Ivezic, Z., Krabbendam, V., Monet, D., O'Connor, P., Oliver, J., Olszewski, E., Pinto, P., Saha, A., Smith, C., Stubbs, C., Takacs, P., & Tyson, A. 2006, in Ground-based and Airborne Telescopes. Edited by Stepp, Larry M.. Proceedings of the SPIE, Volume 6267, pp. (2006).

Coil, A. L., Davis, M., Madgwick, D. S., Newman, J. A., Conselice, C. J., Cooper, M., Ellis, R. S., Faber, S. M., Finkbeiner, D. P., Guhathakurta, P., Kaiser, N., Koo, D. C., Phillips, A. C., Steidel, C. C., Weiner, B. J., Willmer, C. N. A., & Yan, R. 2004a, ApJ, 609, 525

Coil, A. L., Gerke, B. F., Newman, J. A., Ma, C.-P., Yan, R., Cooper, M. C., Davis, M., Faber, S. M., Guhathakurta, P., & Koo, D. C. 2006a, ApJ, 638, 668

Coil, A. L., Hennawi, J. F., Newman, J. A., Cooper, M. C., & Davis, M. 2006b, ArXiv Astrophysics e-prints

Coil, A. L., Newman, J. A., Cooper, M. C., Davis, M., Faber, S. M., Koo, D. C., & Willmer, C. N. A. 2006c, ApJ, 644, 671

Coil, A. L., Newman, J. A., Croton, D., Cooper, M. C., Davis, M., Faber, S. M., Gerke, B. F., Koo, D. C., Padmanabhan, N., Wechsler, R. H., & Weiner, B. J. 2007, ArXiv e-prints, 708

Coil, A. L., Newman, J. A., Kaiser, N., Davis, M., Ma, C.-P., Kocevski, D. D., & Koo, D. C. 2004b, ApJ, 617, 765

Cooper, M. C., Newman, J. A., Croton, D. J., Weiner, B. J., Willmer, C. N. A., Gerke, B. F., Madgwick, D. S., Faber, S. M., Davis, M., Coil, A. L., Finkbeiner, D. P., Guhathakurta, P., & Koo, D. C. 2006, MNRAS, 370, 198

Cooray, A. & Sheth, R. 2002, Phys. Rep., 372, 1

Deustua, S., Curtis, D., Goldhaber, G., Graham, J. R., Harris, S., Harvey, P., Heetderks, H., Kim, A., Lampton, M., Lin, R., Pankow, D., Pennypacker, C., Spadafora, A., Smoot, G. F., Aldering, G., Edwards, W., Frye, B., Groom, D., Holland, S., Kasen, D., Knop, R., Lafever, R., Levi, M., Nugent, P., Perlmutter, S., Robinson, K., Astier, P., Genat, J. F., Hardin, D., Levy, J.-M., Pain, R., Schamahneche, K., Baden, A., Goodman, J., Sullivan, G., Ellis, R., Metzger, M. R., Huterer, D., Fruchter, A., Bebek, C., Bergstrom, L., Goobar, A., Lidman, C., Rich, J., Mourao, A., & SNAP Collaboration. 2000, in Bulletin of the American Astronomical Society, 722–+

Hoekstra, H., Mellier, Y., van Waerbeke, L., Semboloni, E., Fu, L., Hudson, M. J., Parker, L. C., Tereno, I., & Benabed, K. 2006, ApJ, 647, 116

Huterer, D. 2002, Phys. Rev. D, 65, 063001

Huterer, D., Kim, A., Krauss, L. M., & Broderick, T. 2004, ApJ, 615, 595

Huterer, D., Takada, M., Bernstein, G., & Jain, B. 2006, MNRAS, 366, 101

Ilbert, O., Arnouts, S., McCracken, H. J., Bolzonella, M., Bertin, E., Le Fèvre, O., Mellier, Y., Zamorani, G., Pellò, R., Iovino, A., Tresse, L., Le Brun, V., Bottini, D., Garilli, B., Maccagni, D., Picat, J. P., Scaramella, R., Scodeggio, M., Vettolani, G., Zanichelli, A., Adami, C., Bardelli, S., Cappi, A., Charlot, S., Ciliegi, P., Contini, T., Cucciati, O., Foucaud, S., Franzetti, P., Gavignaud, I., Guzzo, L., Marano, B., Marinoni, C., Mazure, A., Meneux, B., Merighi, R., Paltani, S., Pollo, A., Pozzetti, L., Radovich, M., Zucca, E., Bondi, M., Bongiorno, A., Busarello, G., de La Torre, S., Gregorini, L., Lamareille, F., Mathez, G., Merluzzi, P., Ripepi, V., Rizzo, D., & Vergani, D. 2006, A&A, 457, 841

Ilbert, O., Tresse, L., Zucca, E., Bardelli, S., Arnouts, S., Zamorani, G., Pozzetti, L., Bottini, D., Garilli, B., Le Brun, V., Le Fèvre, O., Maccagni, D., Picat, J.-P., Scaramella, R., Scodeggio, M., Vettolani, G., Zanichelli, A., Adami, C., Arnaboldi, M., Bolzonella, M., Cappi, A., Charlot, S., Contini, T., Foucaud, S., Franzetti, P., Gavignaud, I., Guzzo, L., Iovino, A., McCracken, H. J., Marano, B., Marinoni, C., Mathez, G., Mazure, A., Meneux, B., Merighi, R., Paltani, S., Pello, R., Pollo, A., Radovich, M., Bondi, M., Bongiorno, A., Busarello, G., Ciliegi, P., Lamareille, F., Mellier, Y., Merluzzi, P., Ripepi, V., & Rizzo, D. 2005, A&A, 439, 863

Jimenez, R. & Haiman, Z. 2006, Nature, 440, 501

Kaiser, N. 1987, MNRAS, 227, 1
—. 1998, ApJ, 498, 26
Knox, L., Song, Y.-S., & Zhan, H. 2006, ArXiv Astrophysics e-prints
Kochanek, C. S., Eisenstein, D., Caldwell, N., Cool, R., Green, P., & AGES. 2004, in Bulletin of the American Astronomical Society, 1495–+
Le Fèvre, O., Vettolani, G., Garilli, B., Tresse, L., Bottini, D., Le Brun, V., Maccagni, D., Picat, J. P., Scaramella, R., Scodeggio, M., Zanichelli, A., Adami, C., Arnaboldi, M., Arnouts, S., Bardelli, S., Bolzonella, M., Cappi, A., Charlot, S., Ciliegi, P., Contini, T., Foucaud, S., Franzetti, P., Gavignaud, I., Guzzo, L., Ilbert, O., Iovino, A., McCracken, H. J., Marano, B., Marinoni, C., Mathez, G., Mazure, A., Meneux, B., Merighi, R., Paltani, S., Pellò, R., Pollo, A., Pozzetti, L., Radovich, M., Zamorani, G., Zucca, E., Bondi, M., Bongiorno, A., Busarello, G., Lamareille, F., Mellier, Y., Merluzzi, P., Ripepi, V., & Rizzo, D. 2005, A&A, 439, 845
Lilly, S. & The Zcosmos Team. 2005, The Messenger, 121, 42
Ma, Z., Hu, W., & Huterer, D. 2006, ApJ, 636, 21
Masjedi, M., Hogg, D. W., Cool, R. J., Eisenstein, D. J., Blanton, M. R., Zehavi, I., Berlind, A. A., Bell, E. F., Schneider, D. P., Warren, M. S., & Brinkmann, J. 2006, ApJ, 644, 54
Newman, J. A. & Davis, M. 2002, ApJ, 564, 567
Padmanabhan, N., Schlegel, D. J., Seljak, U., Makarov, A., Bahcall, N. A., Blanton, M. R., Brinkmann, J., Eisenstein, D. J., Finkbeiner, D. P., Gunn, J. E., Hogg, D. W., Ivezic, Z., Knapp, G. R., Loveday, J., Lupton, R. H., Nichol, R. C., Schneider, D. P., Strauss, M. A., Tegmark, M., & York, D. G. 2006, ArXiv Astrophysics e-prints
Peebles, P. J. E. 1980, The large-scale structure of the universe (Research supported by the National Science Foundation. Princeton, N.J., Princeton University Press, 1980. 435 p.)
Perlmutter, S. & SNAP. 2004, in Bulletin of the American Astronomical Society, 1559–+
Phillipps, S. 1985, MNRAS, 212, 657
Phillipps, S. & Shanks, T. 1987, MNRAS, 227, 115
Schneider, M., Knox, L., Zhan, H., & Connolly, A. 2006, ArXiv Astrophysics e-prints
Scranton, R., Ménard, B., Richards, G. T., Nichol, R. C., Myers, A. D., Jain, B., Gray, A., Bartelmann, M., Brunner, R. J., Connolly, A. J., Gunn, J. E., Sheth, R. K., Bahcall, N. A., Brinkman, J., Loveday, J., Schneider, D. P., Thakar, A., & York, D. G. 2005, ApJ, 633, 589

Seo, H.-J. & Eisenstein, D. J. 2003, ApJ, 598, 720
Smith, R. E., Peacock, J. A., Jenkins, A., White, S. D. M., Frenk, C. S., Pearce, F. R., Thomas, P. A., Efstathiou, G., & Couchman, H. M. P. 2003, MNRAS, 341, 1311
Spergel, D. N., Bean, R., Dore', O., Nolta, M. R., Bennett, C. L., Hinshaw, G., Jarosik, N., Komatsu, E., Page, L., Peiris, H. V., Verde, L., Barnes, C., Halpern, M., Hill, R. S., Kogut, A., Limon, M., Meyer, S. S., Odegard, N., Tucker, G. S., Weiland, J. L., Wollack, E., & Wright, E. L. 2006, ArXiv Astrophysics e-prints
Steidel, C. C., Adelberger, K. L., Giavalisco, M., Dickinson, M., & Pettini, M. 1999, ApJ, 519, 1
Steidel, C. C., Adelberger, K. L., Shapley, A. E., Pettini, M., Dickinson, M., & Giavalisco, M. 2003, ApJ, 592, 728
Steidel, C. C., Shapley, A. E., Pettini, M., Adelberger, K. L., Erb, D. K., Reddy, N. A., & Hunt, M. P. 2004, ApJ, 604, 534
Strauss, M. A., Weinberg, D. H., Lupton, R. H., Narayanan, V. K., Annis, J., Bernardi, M., Blanton, M., Burles, S., Connolly, A. J., Dalcanton, J., Doi, M., Eisenstein, D., Frieman, J. A., Fukugita, M., Gunn, J. E., Ivezić, Ž., Kent, S., Kim, R. S. J., Knapp, G. R., Kron, R. G., Munn, J. A., Newberg, H. J., Nichol, R. C., Okamura, S., Quinn, T. R., Richmond, M. W., Schlegel, D. J., Shimasaku, K., SubbaRao, M., Szalay, A. S., Vanden Berk, D., Vogeley, M. S., Yanny, B., Yasuda, N., York, D. G., & Zehavi, I. 2002, AJ, 124, 1810
Tyson, A. 2005, in ASP Conf. Ser. 339: Observing Dark Energy, ed. S. C. Wolff & T. R. Lauer, 95–+
Tyson, A. & Angel, R. 2001, in ASP Conf. Ser. 232: The New Era of Wide Field Astronomy, ed. R. Clowes, A. Adamson, & G. Bromage, 347–+
Tyson, J. A. 2006, ArXiv Astrophysics e-prints
Zehavi, I., Zheng, Z., Weinberg, D. H., Frieman, J. A., Berlind, A. A., Blanton, M. R., Scoccimarro, R., Sheth, R. K., Strauss, M. A., Kayo, I., Suto, Y., Fukugita, M., Nakamura, O., Bahcall, N. A., Brinkmann, J., Gunn, J. E., Hennessy, G. S., Ivezić, Ž., Knapp, G. R., Loveday, J., Meiksin, A., Schlegel, D. J., Schneider, D. P., Szapudi, I., Tegmark, M., Vogeley, M. S., & York, D. G. 2005, ApJ, 630, 1
Zhan, H. 2006, Journal of Cosmology and Astro-Particle Physics, 8, 8
Zhan, H. & Knox, L. 2006, ApJ, 644, 663