FERMILAB-PUB-504-AE-CD

The Baryonic Acoustic Feature and Large-Scale Clustering in the SDSS LRG Sample

Eyal A. Kazin,[a1] Michael R. Blanton,[a] Román Scoccimarro,[a] Cameron K. McBride,[b] Andreas A. Berlind,[b] Neta A. Bahcall,[c] Jon Brinkmann,[d] Paul Czarapata,[e] Joshua A. Frieman,[fg] Stephen M. Kent,[h] Donald P. Schneider,[i] Alexander S. Szalay [j]

[a] *Center for Cosmology and Particle Physics, New York University, 4 Washington Place, New York, NY 10003, USA*

[b] *Department of Physics and Astronomy, Vanderbilt University, 1807 Station B, Nashville, TN 37235, USA*

[c] *Department of Astrophysical Sciences, Princeton University, Princeton, NJ 08544, USA*

[d] *Apache Point Observatory, P.O. Box 59, Sunspot, NM 88349, USA*

[e] *Fermi National Accelerator Laboratory, P..O Box 500, Batavia, IL 60510, USA*

[f] *Particle Astrophysics Center, Fermilab, P.O. Box 500, Batavia, IL 60510, USA*

[g] *Kavli Institute for Cosmological Physics, Department of Astronomy & Astrophysics, University of Chicago, Chicago, IL 60637, USA*

[h] *MS 127, Fermilab, P. O. Box 500, Batavia, IL 60510, USA*

[i] *Department of Astronomy and Astrophysics, The Pennsylvania State University, 525 Davey Laboratory, University Park, PA 16802, USA*

[j] *Department of Physics and Astronomy, The Johns Hopkins University, 3701 San Martin Drive, Baltimore, MD 21218, USA*

ABSTRACT

We examine the correlation function ξ of the Sloan Digital Sky Survey (SDSS) Luminous Red Galaxy sample (LRG) at large scales ($60 < s < 400\,h^{-1}$Mpc) using the final data release (DR7). Focusing on a quasi volume-limited ($0.16 < z < 0.36$) subsample and utilizing mock galaxy catalogs, we demonstrate that the observed baryonic acoustic peak and larger scale signal are consistent with ΛCDM at $70-95\%$ confidence. Fitting data to a non-linear, redshift-space, template based-model, we constrain the peak position at $s_{\rm p}=101.7\pm 3.0\,h^{-1}$Mpc when fitting the range $60 < s < 150\,h^{-1}$Mpc (1σ uncertainties). This redshift-space distance $s_{\rm p}$ is related to the comoving sound horizon scale r_s after taking into account matter clustering non-linearities, redshift distortions and galaxy clustering bias. Mock catalogs show that the probability that a DR7-sized sample would not have an identifiable peak is at least $\sim 10\%$. As a consistency check of a fiducial cosmology, we use the observed $s_{\rm p}$ to obtain the distance $D_V \equiv [(1+z)^2 D_A^2 cz/H(z)]^{\frac{1}{3}}$

relative to the acoustic scale. We find $r_s/D_V(z = 0.278) = 0.1389 \pm 0.0043$. This result is in excellent agreement with Percival et al. (2009), who examine roughly the same data set, but use the power spectrum. Comparison with other determinations in the literature are also in very good agreement. The signal of the full sample at $125 < s < 200\,h^{-1}$Mpc tends to be high relative to theoretical expectations; this slight deviation can probably be attributed to sample variance. We have tested our results against a battery of possible systematic effects, finding all effects are smaller than our estimated sample variance.

Subject headings: cosmology: observation, large-scale structure of universe, distance scale, galaxies: elliptical and lenticular, cD

1. Introduction

The large scale clustering of matter is a critical test of any cosmological model. An interesting feature in the auto-correlation-function ξ is the baryonic acoustic peak, a residual from plasma sound-waves that came to a near stop at the end of the baryon drag epoch ($z_d \sim 1010$). The baryonic acoustic signature is strongly imprinted in the Cosmic Microwave Background (CMB) temperature fluctuations and observations determine its physical size to high precision ($< 1.3\%$; Spergel et al. 2003). Predicted to appear in galaxy clustering measurements, the baryonic acoustic feature can in principle provide determination of cosmic distances with very small statistical uncertainties ($< 1\%$; Seo & Eisenstein 2007).

The feature in the galaxy ξ was first measured by Eisenstein et al. (2005) using Luminous Red Galaxies (LRG) from the Sloan Digital Sky Survey (SDSS; York et al. 2000). Using $\sim 47,000$ LRGs from the third data release (DR3), they determined the distance to $z \sim 0.35$ at 5% accuracy and constrained cosmological parameters. Similar measurements using later data releases detected a peak, but as we discuss below, a broader one (Cabré & Gaztañaga 2009; Martinez et al. 2008; Labini et al. 2009; Sanchez et al. 2009). This apparent peculiarity may raise concerns about the utility of the baryonic acoustic peak in distance determination and the consistency of ΛCDM models in large-scale galaxy clustering.

Padmanabhan et al. (2007) (DR3; Abazajian et al. 2005) and Blake et al. (2007) (DR4; Adelman-McCarthy et al. 2006) showed in their photo-z analysis of the power spectrum of the LRGs a hint of excess power at large scales. Blake et al. (2007) suggest a few possible causes for this excess

[1]eyalkazin@gmail.com

power: residual systematic errors, cosmic variance, large-scale galaxy biasing mechanisms, and new early Universe physics. Cabré & Gaztañaga (2009) used Data Release 6 (DR6; Adelman-McCarthy & Sloan Digital Sky Survey 2008; $\sim 75,000$ LRGs, $\sim 1\,h^{-3}\mathrm{Gpc}^3$ in comoving-volume) to probe for possible systematics in data analysis, showing that the strong large-scale signal at $s > 130\,h^{-1}\mathrm{Mpc}$ is persistent for various choices of weighting schemes and galaxy sample selection. Martinez et al. (2008) confirmed this stable, but wider baryonic acoustic feature in DR7 (Abazajian & Sloan Digital Sky Survey 2009). ΛCDM models predict a ξ that crosses over from positive correlations to negative correlations at scales of $r_c \sim 140\,h^{-1}\mathrm{Mpc}$ (Peebles 1980, Gabrielli et al. 2002). Labini et al. (2009) reanalyzed DR7, pointing out that there was no obvious cross-over to anti-correlations at this scale.

Other studies that investigate large-scale clustering and the baryonic acoustic feature of the SDSS LRGs and the 2dF Galaxy Redshift Survey (Colless et al. 2003) include Cole et al. (2005), Tegmark et al. (2006), Hütsi (2006), Percival et al. (2007), Sanchez et al. (2009), Percival et al. (2009), and Reid et al. (2009).

The purpose of this study is to measure the LRG two-point correlation function on large scales, compare it with the predictions of a ΛCDM model, and derive constraints on the baryon acoustic oscillation scale. We obtain precise uncertainty estimates on ξ by using a large suite of mock catalogs drawn from N-body simulations. In the course of our study, we examine the differences in ξ at large scales among the different SDSS data releases; we show that the differences between the DR3 and DR7 (the latter contains $\sim 105,000$ LRGs, $\sim 1.6\,h^{-3}\mathrm{Gpc}^3$) results do not arise from known systematics in the data analysis. Further, focusing on a quasi volume-limited subsample we demonstrate that the stronger large-scale signal in the final data release is consistent with the ΛCDM framework. We also analyze the position of the baryonic acoustic feature, explaining systematics, and relate our measurement to the cosmic distance D_V (Equation 8).

In §2 we discuss the SDSS data set. In §3 we explain our methods and mock catalogs. In §4.1 we show that we can reproduce results from Eisenstein et al. (2005) while sampling a similar DR3 sample. In §4.2 and §4.3 we apply the same technique to calculate the large-scale ξ of DR7 volume limited subsamples and compare to LRG mock catalogs. In §4.2 we also investigate the chances of identifying a baryon acoustic peak in an SDSS-sized sample. In §4.4 we test the robustness of the baryonic acoustic feature and large-scale signal in ξ to systematic errors in calibration, suggesting that SDSS photometric calibration errors should not affect the final results. In §4.5 we determine the location of the peak of the baryonic acoustic feature, and its uncertainty. In §4.6 we use the peak position to determine the ratio between the acoustic scale at the drag epoch r_s to an effective distance D_V at redshift of $z = 0.278$. In the appendices we discuss technical aspects of the selection function, systematic uncertainties and correlation estimators.

In the following, unless otherwise indicated, all calculations assume a flat ΛCDM model with

present day matter density $\Omega_{M0} = 0.25$ and a present Hubble expansion rate $H_0 = 100h$ km s^{-1} Mpc^{-1}. When running mock simulations $h = 0.7$, but otherwise $h = 1$ when converting redshifts to comoving distances. When analyzing the DR3 we use the same cosmology as Eisenstein et al. (2005), $\Omega_{M0} = 0.3$.

2. Data

The SDSS is the largest volume LRG survey to date, having imaged the sky at high Galactic latitude in five passbands u, g, r, i and z (Fukugita et al. 1996, Gunn et al. 1998) using a 2.5m telescope (Gunn et al. 2006). The images are processed (Lupton et al. 2001, Stoughton et al. 2002, Pier et al. 2003, Ivezić et al. 2004) and calibrated (Hogg et al. 2001, Smith et al. 2002, Tucker et al. 2006), allowing selection of galaxies, quasars (Richards et al. 2002) and stars for spectroscopy (Eisenstein et al. 2001; Strauss et al. 2002) with twin fiber fed double spectographs. Targets are assigned to plug plates according to a tiling algorithm ensuring nearly complete samples (Blanton et al. 2003).

The SDSS LRG sample developed by Eisenstein et al. (2001) serves as a good tracer of matter as they are associated with massive halos. The LRG's high luminosity enables us to obtain a large volume, and their spectral uniformity make them relatively easy to identify.

The SDSS LRG sample covers 19% of the sky (distribution is shown in Figure 1). The sample includes a large quasi-volume-limited region to a redshift of $z \sim 0.36$, and is flux-limited thereafter extending to $z \sim 0.47$. The peak at $z \sim 0.36$ is associated with the passage of the 4000Å break into the r-band.

From the full sample DR7-Full ($0.16 < z < 0.47$, $-23.2 < M_g < -21.2$) we subsample for two main purposes: comparison with previous studies, focus on volume-limited regions.

To avoid the effects of the flux limited region, we focus much of our analysis on subsamples of DR7 similar to those chosen by Zehavi et al. (2005). DR7-Dim ($0.16 < z < 0.36$, $-23.2 < M_g < -21.2$) is produced by subsampling DR7-Full at $z \leqslant 0.36$. This quasi-volume-limited subsample is *not* a dimmer sample than DR7-Full, but rather is called "Dim" as a moniker to distinguish from the overlapping brighter subsample DR7-Bright ($0.16 < z < 0.44$, $-23.2 < M_g < -22.8$). We also include results for another volume limited subsample DR7-Bright2, which is a brighter subsample of DR7-Dim ($0.16 < z < 0.36$, $-22.6 < M_g < -21.6$).

All subsamples are summarized in Table 1. In addition to various cuts in redshift and absolute magnitude M_g, we also analyze a sample limited to the earlier release DR3 analyzed by Eisenstein et al. (2005).

Figure 2 shows the comoving number density as a function of redshift $n(z)$ for each sample. DR7-Full (black) is the full sample, from which we subsample DR3 (cyan). DR3 covers only the sky area of DR7-Full which was also covered by DR3. For most calculations in DR3 we apply the fiducial cosmology chosen by Eisenstein et al. (2005). It yields a similar (but noisier) $n(z)$ and comoving volume density to the DR7-Full, differing somewhat due to the different cosmologies (flat ΛCDM, with $\Omega_{M0} = 0.25$ in DR7 and $\Omega_{M0} = 0.3$ in DR3). Our subsample agrees in detail with Eisenstein et al. (2005) with only < 350 mismatches out of $\sim 47,000$ galaxies (0.7%). For a more thorough discussion of LRG selection please see Appendix A.1. The figure also displays the radial selection function $n(z)$ of subsamples DR7-Dim (green), DR7-Bright (blue) and DR7-Bright2 (red).

In Appendix A.2 we show, using mock catalogs, that features in $n(z)$ up to $z < 0.36$ cause a negligible change in the ξ noise properties of the sample relative to a volume-limited one.

For reasons described in §3.2 we only use the Northern Galactic Cap for subsamples DR7-Dim and DR7-Bright. For DR7-Full we make use of both caps, as in DR3. As explained in Appendix B.3, we verified that the resulting ξ does not change when excluding the Southern Galactic Cap region.

A physical constraint when obtaining spectra, known as fiber-collisions, is dealt with here when analyzing the data. Due to the physical size of the fibers of the spectrometer, one can not obtain spectra simultaneously of two targeted objects that reside within $55''$ of each other. We count $\sim 2\%$ of targeted LRGs were missed due to fiber-collisions. To account for this we up-weight LRGs with spectra which are within $55''$ proximity of LRGs without. For more details of our method for fiber-collision correction, its impact on ξ, as well as accounting for holes and boundary effects of the survey, please refer to Appendix A.1.

All data used in this study may be obtained on the World Wide Web.[2]

3. Method

3.1. Clustering Estimator and Random Points

To measure the two-point correlation function we use the Landy & Szalay (1993) estimator:

$$\xi(s) = \frac{DD(s) - 2DR(s) + RR(s)}{RR(s)}, \tag{1}$$

[2]http://cosmo.nyu.edu/∼eak306/SDSS-LRG.html

which compares the normalized number of data pair counts to randomly distributed points. The quantity s refers to the mean redshift space separation for each bin. If we define r to be the ratio of the number of random points to data and $N_{DD}(s)$ to be the total number of galaxy pairs separated by values $(s - ds/2, s + ds/2]$ (where ds is the width of the bin), then the normalized number of pairs are $DD = N_{DD}r^2$, $DR = N_{DR}r$, and $RR = N_{RR}$. Here, DR and RR stand for data-random and random-random pairs, respectively. The random points account for the effects introduced by survey boundary, holes within the data set, and sector incompleteness.

Additionally, in our counting of pairs, we apply weights to each galaxy. Appendices A and B.2 discuss the details of distributing the random points and the pair-count weighting. In Appendix B.3, we also compare this estimator to other known methods, showing excellent agreement with the method proposed by Hamilton (1993) on large scales, and substantial differences with those proposed by Davis & Peebles (1983) and Peebles & Hauser (1974).

3.2. LRG Mock Catalogs

We use mock galaxy catalogs produced from the LasDamas simulations (McBride et al. 2009; in prep) to measure the uncertainty covariance matrix, as well as to investigate systematic errors in the two-point-correlation estimators. These mock catalogs provide 160 light-cone redshift-space realizations of an SDSS-sized volume, with appropriate number densities and clustering properties for comparison to the observed data.

The LasDamas simulations are designed to model the clustering of SDSS DR7 in a wide luminosity range. In this suite of simulations galaxies are artificially placed in dark matter halos specifically using the formalism of the halo occupation distribution (HOD; Berlind & Weinberg 2002) with parameters chosen to match an observational SDSS sample. For complete details see McBride et al. (2009, in prep.); for distribution visit the World Wide Web.[3]

We use the "gamma" release of mock LRG catalogs produced from simulations. The Oriana simulation consists of 40 N-body dark matter simulations; each realization contains 1280^3 particles of mass $45.73 \times 10^{10} h^{-1}\mathrm{M}_\odot$ in a box of length $2400\, h^{-1}$Mpc. The softening parameter is $53h^{-1}$kpc. The simulations assume a flat ΛCDM cosmology with total matter density $\Omega_{M0} = 0.25$, baryon density $\Omega_{b0} = 0.04$, Hubble expansion rate $H_0 = 70\,\mathrm{km\ s^{-1}Mpc^{-1}}$, $\sigma_8 = 0.8$, and spectral index $n_s = 1$. The HOD parameters are adjusted to reproduce the observed number density as well as the projected two-point-correlation function $w_p(r_p)$ at projected separations $0.3 < r_p < 30\, h^{-1}$Mpc, well below the scales we consider here. The choice of HOD affects the resulting $\xi(s)$ shape on small

[3]http://lss.phy.vanderbilt.edu/lasdamas/

scales, but on large scales the HOD choice primarily affects the amplitude as expected from local galaxy bias arguments (Fry & Gaztanaga 1993; Coles 1993; Scherrer & Weinberg 1998; Narayanan et al. 2000).

The LasDamas team has divided the sky into two alternative footprints: one can use either the SDSS Northern Galactic and Southern Galactic Caps together, or only the Northern. We use the latter option because in that case each simulation produces four samples (as opposed to two in the former) resulting in twice as many realizations. In summary, we analyze $40 \times 4 = 160$ LasDamas mock catalogs for each of the luminosity subsamples DR7-Bright and DR7-Dim. We note that the angular distribution of galaxies in these mocks is similar to the angular distribution of the data, aside from the observational incompleteness. In Appendix A.1 we explain how we account for the incompleteness within the observational data.

Our only manipulation of the catalogs provided by McBride et al. (2009; in prep) is in the radial direction, where we randomly subsample to match the SDSS comoving volume density $n(z)$, to better represent the Poisson noise. This subsampling reduces the number of LRGs in DR7-Dim by 15% and in DR7-Bright by 7%. In Appendix A.2 we show that this subsampling does not noticeably affect either ξ or its uncertainties. For more details regarding the radial selection function, please refer to that Appendix.

3.3. Covariance Matrix and χ^2 Fitting

In our analysis we use the standard χ^2 fitting algorithm based on a covariance matrix, which is constructed from the LasDamas mock realizations.

Because the uncertainties in the ξ values for the bins are interdependent, we build a covariance matrix C_{ij} to measure the dependence of the ith bin on the jth. We construct C_{ij} from the individual mock realizations as follows:

$$C_{ij} = \frac{1}{N_{\rm mocks} - 1} \cdot \sum_{k=1}^{N_{\rm mocks}} \left(\overline{\xi}_i - \xi_i^k\right) \left(\overline{\xi}_j - \xi_j^k\right), \tag{2}$$

where $\overline{\xi}_m$ is the correlation value for the mth bin of the mock mean and ξ_m^k is the same for the kth mock realization. In all calculations here, $N_{\rm mocks} = 160$.

We can then estimate χ^2 for our observational result $\xi^{\rm obs}$ relative to models $\xi^{\rm model}$ using:

$$\chi^2(\vec{\theta}) = \sum_{i,j=1}^{N_s} \left(\xi_i^{\rm obs} - \xi_i^{\rm model}(\vec{\theta})\right) C_{ij}^{-1} \left(\xi_j^{\rm obs} - \xi_j^{\rm model}(\vec{\theta})\right), \tag{3}$$

where $\vec{\theta}$ are parameters of the model, and N_s is the number of separation bins used.

4. Results

Here we present our results and analysis of the redshift space angle-averaged $\xi(s)$ of the SDSS LRGs in two releases: DR3 and DR7. In §4.1 we show that we can reproduce the DR3 LRG selection as well as the $\xi(s)$ results of Eisenstein et al. (2005). Differences between DR3 and DR7 are, therefore, not due to systematic differences in our analysis from that of Eisenstein et al. (2005). In §4.2 we use mock catalogs to show that the large scale clustering of DR7 LRGs is consistent with the concordance ΛCDM model at the $70-95\%$ confidence level. Furthermore, in §4.3, we investigate a bright subsample, which has a strong large-scale signal relative to our ΛCDM predictions, not seen in the other subsamples. In §4.4, we investigate a possible observational bias, and show the robustness of the baryonic acoustic feature and large-scale clustering to this effect. In §4.5 we measure the peak position of the baryonic acoustic feature. We use this last measurement in §4.6 to determine the ratio r_s/D_V and compare results to other studies.

4.1. Reproducing the DR3 $\xi(s)$

To demonstrate consistency with previous studies, we first compare our analysis of an earlier data release, DR3, with that of Eisenstein et al. (2005). For this analysis we create a DR3 LRG sample with the same criteria used by Eisenstein et al. (2005), and a corresponding random sample. In this subsection we assume $\Omega_{M0} = 0.3$ in calculations.

In Figure 3 we reproduce the redshift-space ξ of the SDSS LRGs first measured by Eisenstein et al. (2005). There is good agreement between our results (green diamonds) and theirs (red crosses). To investigate the minute remaining differences, we also test using their random catalogs with our data set and vice-versa. We conclude that the remaining subtle variations are due to Poisson noise in the random catalogs. As previously presented in Eisenstein et al. (2005), we obtain a narrow peak at an apparent separation $s > 100\,h^{-1}$Mpc (see §4.5 for our analysis of peak position value), and a steep slope that flattens at $\sim 135\,h^{-1}$Mpc.

The dashed lines in Figure 3 display our result for the final data release DR7 (sample DR7-Full), which runs through the same galaxy selection algorithm and data analysis as DR3. The thick blue dashed line uses the same cosmology as in DR3; for the thin black dashed line we alter that cosmology to $\Omega_{M0} = 0.25$, $\Omega_{\Lambda 0} = 0.75$. The binning in all cases is the same chosen by Eisenstein et al. (2005), with the exception that for DR7 we extend the signal a bit further. In both DR7 cases the resulting position of the baryonic acoustic feature is in fair agreement (§4.5). However, we clearly see stronger power on large scales, yielding a wider peak. In the next section we examine the significance of this strong large-scale signal relative to our ΛCDM model predictions.

4.2. Consistency of SDSS LRGs Clustering with ΛCDM

We test the DR7 LRG clustering against a ΛCDM model by comparing the observed ξ to those yielded by the LasDamas mock LRG catalogs (§3.2). We run the same analysis as before on each subsample of DR7: DR7-Dim and DR7-Bright (see Table 1 and Figure 2). We also analyze DR7-Full, which has the disadvantage of being flux limited at $z > 0.36$ but can probe larger-scale modes. We find that DR7-Dim is in clear agreement with the model used in the simulations. DR7-Bright also has a strong signal at $s > 150\,h^{-1}$Mpc.

We first investigate the quasi-volume limited subsample DR7-Dim showing the SDSS results compared to the 160 LasDamas mock realizations. This subsample has a similar $\xi(s)$ to DR7-Full at most radii $s < 145\,h^{-1}$Mpc (see Figure 4 for direct comparison). Our results are presented in Figure 5. The mock mean is indicated by the green solid line and the SDSS results are the diamonds. The light gray shaded region represents the area containing 68.2% of the realizations lying closest to the mean (that is, the 1σ uncertainties). The dark gray area is the same for 95.4% (2σ uncertainties). The dotted blue lines indicate the strongest and weakest signal of all mocks at each separation (not one realization in particular).

The observed signal is clearly within the 1–2σ significance level at all scales. To quantify the significance of the strong signal at large scales ($s > 130\,h^{-1}$Mpc) we reproduce this last result to $s < 400\,h^{-1}$Mpc (top inset; larger separation binning). Although the expected signal at the largest scales is very small relative to the noise, in the top inset we include data as far out as possible without significantly reducing pair counts due to edge effects. For a discussion regarding edge effects, please refer to Appendix B.3.

We also show in Figure 5 a histogram of the mean correlation on large scales $\langle\xi\rangle_{s=[130,400]h^{-1}\mathrm{Mpc}}$ (bottom inset; red dashed line is SDSS, green dot-dashed is the mock mean), where the brackets denote an average over all separation bins in the indicated range. To further quantify the significance of the large-scale signal, we measure the χ^2 difference between the observed ξ and the mock mean using Equations 2 and 3. In the last equation we limit our bins to those between $130 < s < 400\,h^{-1}$Mpc. Using the 10 bins in the top inset (d.o.f= 10) we measure a normalized value of $\chi^2/\mathrm{d.o.f} = 0.721$. We have tried several definitions for the significance of large-scale power and all agree that there is a 1.5σ consistency with respect to the ΛCDM plus HOD model used here.

Figure 6 shows several hand-picked mock realizations to demonstrate the effects of variance. We caution the reader to avoid interpreting these particular realizations as "typical". Instead, we have chosen them to demonstrate the variety of signals that could be reasonably expected from a survey the size of SDSS DR7-Dim. Whereas some realizations (e.g., green-dotted and red-solid lines) have a similar clustering signal to that observed (symbols), others do not show evidence of the baryonic acoustic feature (e.g. the orange dashed and cyan dot-dashed lines).

Visually examining at all 160 DR7-Dim realizations, we find that at least 17 ($\sim 10\%$) of the mocks have no peak at the expected $s \sim 100\,h^{-1}$Mpc. We examine ξ results and define a peak-less realization as one with no sign of a peak within $95 < s < 120\,h^{-1}$Mpc. We took a liberal approach, so this result should be considered a lower bound; i.e realizations with subtle peaks were counted as having a peak. If we ask how many of the mocks have very clear signs of a peak, we find about 75 out of the 160 do. This designation is subjective, of course. It is also difficult to compare these numbers with theoretical expectations, e.g. using random Gaussian field statistics is not enough as weak peaks in the Gaussian case (corresponding to the linear density field) are washed out by nonlinear motions.

To verify the fraction of feature-less realizations, we also checked the mock realizations in another independent mock LRG catalog. In the Horizon-Run mock catalog (Kim et al. 2009), LRG positions are determined by identifying physically self bound dark matter sub-halos that are not tidally disrupted by larger structures. After adjusting their mock galaxy catalogs to fit the SDSS radial function $n(z)$, we ran the same analysis as for LasDamas and found that, of their 32 realizations, 5 showed no sign of the baryonic acoustic feature, comparable to the result obtained with LasDamas mock catalogs.

4.3. Bright Subsample DR7-Bright

Figure 4 also shows that the correlation function of DR7-Full (dashed line) is stronger than DR7-Dim (green diamonds) at scales of $150 < s < 200\,h^{-1}$Mpc. To help understand the difference between these samples we examine another subsample, DR7-Bright.

The correlation function ξ of the brighter volume limited sample, DR7-Bright (blue crosses), is stronger on all scales than DR7-Full and DR7-Dim, which is expected, as bias is known to be a function of luminosity (see e.g. Zehavi et al. 2005). For example, Figure 7 shows the relative redshift-space bias ratios, defined as

$$b \equiv \sqrt{\xi_{DR7-Bright}/\xi_{DR7-Dim}}, \tag{4}$$

showing that DR7-Bright is biased by a factor of 1.14 on most scales relative to DR7-Dim, in agreement with the mock catalogs (by design, of course, at scales $r_p < 30\ h^{-1}$Mpc). The gray bands show one and two σ distributions of the mock results when relating each of the DR7-Bright samples to the same realization DR7-Dim mock sample. The dotted lines are the outermost values for each bin. According to these uncertainties, the relative bias of DR7-Bright is significantly stronger (by at least 2σ) than DR7-Dim up to $s \sim 40\,h^{-1}$Mpc.

As discussed in §4.5, Figure 4 shows good agreement in the baryonic acoustic peak position

among the DR7-Full and DR7-Dim. However, the relative strengths of the large-scale signal is worth investigating. Figure 8 is similar to Figure 5 for DR7-Bright. The figure and its insets show that its signal is significantly stronger than the mock values on scales $s > 130\,h^{-1}$Mpc.

The bottom inset of Figure 8 is the histogram of $\langle\xi\rangle_{s=[130,500]h^{-1}\mathrm{Mpc}}$ for the mocks, compared to the SDSS value (red vertical dashed line). We perform the same χ^2 comparison as in DR7-Dim, but this time for scales up to $500\,h^{-1}$Mpc (see Appendix B.3 for a justification of the larger scale). This test yields $\chi^2/d.o.f = 2.5$ for 11 degrees of freedom (separation bins of top inset), showing an unlikely fit with the model used here.

We note that this result does not rule out all possible ΛCDM and HOD models, because we have compared to only one choice of parameters designed to fit statistics on smaller scales (see §3.2 for details).

We also split DR7-Bright into two subsamples (as indicated in Table 1): DR7-Bright-Near ($0.16 < z < 0.36$; $\sim 16,500$ LRGs) and DR7-Bright-Far ($0.36 < z < 0.44$; $\sim 13,800$ LRGs). Each produce noisy results for ξ, so we can not draw concrete conclusions regarding large-scale clustering and the baryonic acoustic feature. For example, we find, as shown in Cabré & Gaztañaga (2009), that the distant subsample DR7-Bright-Far has no clear sign of a baryonic acoustic feature.

To examine the effect of the luminosity cuts on the strong signal in DR7-Dim in the range $120 < s < 155\,h^{-1}$Mpc, we exclude the DR7-Bright-Near subsample from DR7-Dim by creating in the same volume ($0.16 < z < 0.36$) a subsample including only galaxies dimmer than $M_g = -21.8$. This dimmer subsample ($45,426$ LRGs, $-21.8 < M_g < -21.2$) and DR7-Bright-Near yield low S/N results, so we cannot conclude significant differences.

In summary, we find that the apparent strong signal is consistent with sample variance in all cases. We conclude that, given the current measurement uncertainties, effects of potential systematics, such as different sample redshifts, volumes or intrinsic brightness of the galaxy tracers, are subdominant and not statistically significant.

4.4. Effects of Systematic Calibration Errors

Given the large scales and small signals we are probing here, we need to test whether our results are sensitive to systematic errors in the calibrations as a function of angle. Eisenstein et al. (2001) cautions that the number count of LRG targets is sensitive to small variations in color cuts, especially in the g and i bands. Since the SDSS targeting catalog is known to have calibration errors at the 1% level, the true color cuts applied vary across the survey (Padmanabhan et al. 2008). Such a variation might introduce an artificial clustering signal in our analysis.

To test the possible effect of these errors on ξ, we subsample the LasDamas DR7-Bright sample in a spatially varying fashion, and analyze the correlation function of the resulting sample (using the original, un-subsampled random catalogs). In detail, the subsampling factor varies from 95–100% in sinusoidal waves along the declination direction with a wavelength of 10 deg. This choice is motivated by the fact that the targeting catalog is separately calibrated in each stripe, which spans 2.5 deg and are (very roughly) parallel to lines of constant declination.

Comparing the resulting ξ with the full mock catalogs, we find insignificant effects for individual mock realizations and no difference when averaging over all 160. We tested this both on the full DR7-Bright and on a subsample DR7-Bright-Far ($0.36 < z < 0.44$; see Table 1). Effects are significant only when enhancing the subsampling factor amplitude from 5% to 15%, well above the expected systematic uncertainties due to calibration. Even at this unrealistic uncertainty, the baryonic acoustic feature is noticeable, on average, though with a weaker amplitude.

Notably, we cannot obtain the strong large-scale signal as observed in Figure 8 using any realistic level of calibration uncertainty.

4.5. Baryonic Acoustic Peak Position

Here we measure the baryonic acoustic peak position $s_{\rm p}$. We first construct a model:

$$\xi^{model}(s) = \beta \cdot \overline{\xi}(\alpha \cdot s), \tag{5}$$

where the template $\overline{\xi}$ is based on the mock mean, β represents a bias term and α represents a change in length scale. We minimize χ^2 over α and β (see Equations 2 and 3), using the observed ξ at scales near the peak. After β and α are determined we use a spline interpolation on the (very smooth) best-fit model to pinpoint $s_{\rm p}$. This procedure is run on the observed ξ and, as explained below, for the individual mock realizations.

The advantage of using this approach over linear models is that clustering nonlinearities, clustering bias and redshift distortions are already taken into account in the simulations, as well as the angular mask and radial selection. A disadvantage is that we are not self-consistently altering the cosmology for each model. However, in §4.6, when using $s_{\rm p}$ to determine cosmological distances, we expect the uncertainties in assumptions we make to be small relative to the data uncertainties.

Figure 9 displays our result for DR7-Dim using $\overline{\xi} = \overline{\xi}_{160}$, the mock mean over all 160 realizations. The upper left panel shows the distribution of the best-fit peak position for the mocks and the observations. The vertical red line is the best-fit for the observations, the vertical green line is for the mean mock catalog. The upper right panel exhibits the distribution of χ^2 per degree of freedom. The vertical red line marks the results from the observations and the histogram is that of

75 clear-peaked mocks. The bottom left panel is the normalized covariance matrix $C_{ij}/\sqrt{C_{ii}C_{jj}}$. The bottom right panel compares the best-fit model ξ^{model} (blue solid line) to the observations (symbols), and the red arrow points to the fit peak value.

The mean peak value for the mocks (green line on the top left panel) is $s_{\rm peak}^{\rm mock} = 106.31 \pm 3.19$ h^{-1}Mpc, where the 1σ uncertainties (68.2% C.L) are calculated from mock realizations only with clear signs of a peak (about 75 out of the 160). The individual mock $s_{\rm p}$ results are shown as the histogram on the upper left panel and their normalized χ^2 result in the upper right plot. We obtain the uncertainty on the observations by scaling these uncertainties for this mock catalog distribution by $s_{\rm p}/s_{\rm p}^{mock}$.

Using the observed DR7-Dim sample between $60 < s < 150\,h^{-1}$Mpc we find the peak to be at $s_{\rm p}$= $101.74 \pm 3.05\,h^{-1}$Mpc (3%) where these 1σ uncertainties are scaled from our mock catalog results. We caution that these uncertainties rely on the chosen fiducial cosmology of our mocks and can potentially vary – although we expect such variations to be small within observationally motivated ΛCDM models (see Appendix B.3 for details).

For this fit, we find χ^2/d.o.f. = 1.26, where we use d.o.f.= 19 (the number of data points used minus two fitting parameters β, α). Thus, the fit is within expectations. The bottom right panel of Figure 9 shows a best fit that undershoots the data at all points. This *appears* to be a poor fit only because one's eye does not account for the strong covariances among the bins (see C_{ij} in lower left panel of Figure 9).

If we limit ourselves to the region $60 < s < 135\,h^{-1}$Mpc we find $s_{\rm p} = 105.0 \pm 3.1\,h^{-1}$Mpc. Although both results are consistent, the dependence of $s_{\rm p}$ on fitting range indicates that this sample is still limited in its power to constrain $s_{\rm p}$. We also changed the lower limit to various values between $[55, 75]\,h^{-1}$Mpc, constraining the upper bound to $150\,h^{-1}$Mpc, and find s_p does not vary more than $1.35\,h^{-1}$Mpc. If the lower limit is increased to $80\,h^{-1}$Mpc we obtain $s_{\rm p}$ = $100.7 \pm 3.3\,h^{-1}$Mpc, and increasing to $85\,h^{-1}$Mpc our mechanism for detecting of the peak position breaks down, due to not using the full dip feature around $80\,h^{-1}$Mpc.

We perform the same analysis on 24 jackknife subsamples of the observed sample. They are obtained by dividing the RA-Dec map into $\rm N_{JK} = 24$ regions with equal number of LRGs and excluding one region, each in turn. For each jackknife subsample we calculate ξ and run it through the peak finder algorithm to obtain s_{p_i}. The jackknife uncertainty σ_{JK} is obtained by:

$$\sigma_{JK}^2 = \frac{\rm N_{JK} - 1}{\rm N_{JK}} \sum_{i=1}^{\rm N_{JK}} \left(s_{p_i} - \overline{s}_p^{JK}\right)^2 , \tag{6}$$

where $\overline{s}_p^{JK}$ is the jackknife mean. We measure $\sigma_{JK} = 3.3\,h^{-1}$Mpc. As a test, we jackknife a random (but with an apparent strong baryonic acoustic feature) mock realization, and obtain similar results.

A more in depth analysis of jackknife uncertainties would include testing for different values N_{JK} and geometries for multiple realizations. Our simple test shows jackknife uncertainties of s_p similar to that obtained by sample variance of the mocks. We trust the sample variance of s_p estimated from mock catalogs, as this is obtained by measuring a suite of realistic mocks of the full DR7-Dim volume.

Another template we test is $\overline{\xi} = \overline{\xi}_{75}$, the mock mean of 75 handpicked realizations with a clear peak. $\overline{\xi}_{160}$, used before, yields the expected peak given a DR7-Dim size sample, where $\overline{\xi}_{75}$ produces the same given a realization *with a clear peak* (as in the case of the observations). In Figure 10 we show the correlation functions of $\overline{\xi}_{75}$ (dashed bright green) and $\overline{\xi}_{160}$ (solid black). As expected, both have roughly the same clustering at all scales, except that the latter has a weaker peak. The position of the baryonic acoustic feature is nearly the same ($107.01\,h^{-1}$Mpc for $\overline{\xi}_{75}$ versus $106.31\,h^{-1}$Mpc for $\overline{\xi}_{160}$), and they have roughly the same width. The bottom panel shows the residual $\overline{\xi}_{75} - \overline{\xi}_{160}$. When fitting each template to the data we use the same covariance matrix constructed from all 160 realizations.

When using $\overline{\xi}_{75}$ we obtain for the data between $60 < s < 150\,h^{-1}$Mpc $s_p = 102.03 \pm 2.77\,h^{-1}$Mpc, very similar to that obtained when using $\overline{\xi}_{160}$. The χ^2/d.o.f is 1.21. This shows that the choice of template $\overline{\xi}_{160}$ or $\overline{\xi}_{75}$ does not effect the end result in a DR7-Dim size sample.

Besides sample variance, another source of uncertainty in s_p is due to random-shot noise. In Appendix B.1 we show that choices of different random catalogs (ratio of $r \sim 15.6$ random points per data) yield slightly different s_p. To reduce this effect we choose for our final results to use $r \sim 50$.

We also test the effect of dilution on s_p, i.e., choosing a different cosmology when converting redshifts to comoving distances (Padmanabhan & White 2008). In all above results we apply a flat cosmology with $\Omega_{M0} = 0.25$. When using a different value $\Omega_{M0} = 0.30$ (but maintaining mocks as before) we obtain $s_p = 100.40\,h^{-1}$Mpc at χ^2/d.o.f.$= 1.05$. As this result is well within the 1σ variance and systematics explained above, we conclude that the choice of cosmology does not change our results, but may need to be considered when variance is reduced in future surveys.

We find that the $\overline{\xi}$ for the DR7-Bright subsample, in the same binning used for DR7-Dim, is quite noisy and not useful to measure s_p. Although DR7-Bright covers a larger volume than DR7-Dim, its density is over four times lower yielding a sample with less than half the number of galaxies, severely increasing noise in our measurement.

We applied the same algorithm (same DR7-Dim covariance matrix) on DR7-Full and DR3 and find consistency with these results. A more in-depth analysis would involve building a new covariance matrix for DR7-Full, as it has a larger volume than DR7-Dim. We do not perform this analysis, because the LasDamas mocks do not extend to that volume. DR3 has fewer LRGs than

DR7-Dim resulting in a noisier signal, so we do not find reason to do further analysis.

4.6. Relating Peak Position $s_{\rm p}$ to Cosmic Distance D_V

As Bashinsky & Seljak (2004) describe, the peak position $s_{\rm p}$ is closely related to r_s, the physical comoving sound-wave radius at the end of the baryon drag epoch. r_s is determined to high precision in physical units from CMB measurements (e.g., Komatsu et al. 2009), whereas we determine $s_{\rm p}$ in redshift space. Therefore, we can compare the two to measure the relationship between redshift and distance, as proposed by Eisenstein et al. (1999).

There are three important effects we must consider in such a determination. First, the relationship between the redshift space peak position $s_{\rm p}$ and r_s is sensitive to non-linear clustering, redshift distortions and galaxy bias, the dominant effect being a broadening of the peak (Eisenstein et al. 2007b; Crocce & Scoccimarro 2008; Smith et al. 2008; Angulo et al. 2008; Seo et al. 2008; Eisenstein et al. 2007a; Kim et al. 2009) and a dominant shift towards small scales that is below current statistical uncertainties.

Second, the measured $s_{\rm p}$ is necessarily associated with a fiducial cosmology used to interpret the redshifts and angular positions in terms of comoving distances. The fiducial cosmology we assume is the same concordance flat cosmology as the mocks used ($[\Omega_{M0}^{fid}, \Omega_{b0}^{fid}, h^{fid}] = [0.25, 0.04, 0.7]$). Technically this means that we use cosmological assumptions in two steps of our algorithm: selection of LRGs by magnitude cuts, and converting observed redshifts to comoving distances. For more details please refer to Appendix B.3

Third, we need to specify what "distance" we seek to measure, since the definition of cosmological distance within the context of general relativity is not unique. The relevant distance measures in this context are the angular diameter distance $D_A(z)$ and the Hubble constant $H(z)$ at redshift z (Hogg 1999). The former would be ideally constrained by measuring $s_{\rm p}$ in a thin shell at radius z, and $H(z)$ by measuring the line-of-sight clustering (Matsubara 2004). These measurements individually are hard to perform with the current data set (for attempts in DR6 please refer to Okumura et al. 2008, Gaztañaga et al. 2009). Instead, we will constrain a combination of the two following the standard approach in the literature (see Equation 8 below).

To account for the above effects, the standard interpretation of the acoustic peak in the correlation function is as follows (with an analogous argument made for the power spectrum; Percival et al. 2007). First we assume the proportionality:

$$r_s = \gamma s_p. \tag{7}$$

Mock catalogs can be used to determine γ, as long as the assumed cosmology is not far from the

true one. We do so here for the LasDamas mock catalogs, whose cosmological parameters are well-motivated from the CMB and other constraints. We calculate $r_s^{fid}(z_d) = 159.75$ Mpc for the mocks using Equation 1 from Blake & Glazebrook (2003), and calculate the sound speed c_s and the end of the drag epoch $z_d = 1012.13$ using Equations 4 and 5 from Eisenstein & Hu (1998). We insert $\Theta_{2.7} = 2.725/2.7$ as the temperature ratio of the CMB in their Equation 5. Using s_p^{mock} from §4.5 and $h = 0.7$ we obtain $\gamma = 1.052$.

Second, we construct the "distance" quantity:

$$D_V \equiv \left[(1+z)^2 D_A^2 cz/H(z)\right]^{\frac{1}{3}} . \tag{8}$$

This quantity is designed in such a way that the ratio $D_V(\langle z\rangle)/s_p(\langle z\rangle)$ is approximately independent of the choice of the fiducial cosmological model (Eisenstein et al. 2005; Percival et al. 2007; Padmanabhan & White 2008).

Third, we can rearrange these relationships to obtain:

$$\frac{D_V(\langle z\rangle)}{r_s} = \frac{D_V(\langle z\rangle, \mathrm{fid})}{\gamma s_p(\langle z\rangle)} \tag{9}$$

where $s_p(\langle z\rangle)$ is understood to be the inferred s_p from §4.5 given the fiducial cosmology. Measuring s_p thus yields the ratio of the distance at redshift $\langle z\rangle$ to the acoustic scale. Given the acoustic scale r_s in Mpc from the CMB, we then can determine the distance $D_V(\langle z\rangle)$ to that redshift. As Equation 8 shows, this distance measurement constrains a combination of the angular diameter distance D_A and the Hubble expansion rate $H(\langle z\rangle)$.

Given the results in §4.5 for DR7-Dim, we find $r_s/D_V = 0.1389 \pm 0.0043$ at $\langle z\rangle = 0.278$, with 1σ uncertainties of around 3%. This value is in excellent agreement with the analysis of the power spectrum of a similar sample by Percival et al. (2009). Combining our result with the sound-horizon $r_s = 153.3$ Mpc obtained from WMAP5 CMB (Komatsu et al. 2009) we find $D_V(0.278) = 1103 \pm 43$ Mpc.

Figure 11 presents our result compared with those obtained by other studies, as well as predictions of flat ΛCDM cosmologies. Our data point is the black diamond where we choose to use the mean sample redshift $\langle z\rangle = 0.278$. The red and orange crosses are the values published in Percival et al. (2009) and Percival et al. (2007): $r_s/D_V(z = 0.2) = 0.1981 \pm 0.0071$, $r_s/D_V(0.275) = 0.1390 \pm 0.0037$, $r_s/D_V(0.35) = 0.1094 \pm 0.0040$ where the small red crosses results are indicated in Table 3 of Percival et al. (2009). The cyan cross is $r_s/D_V(0.35) = 0.1097 \pm 0.0039$ obtained by Reid et al. (2009) when analyzing the P(k) of the reconstructed halo density field of DR7 LRGs. The blue square is the result obtained by Eisenstein et al. (2005) ($D_V(0.35) = 1370 \pm 64$ Mpc, if we use the CMB r_s value as before). We also add the result obtained by Sanchez et al. (2009) (purple triangle) who analyze the DR7 LRG ξ. Using only LRG clustering they obtain

$D_V(0.35) = 1230 \pm 220$ Mpc. They show that when combining ξ with CMB measurements they obtain a tighter constraint $D_V(0.35) = 1300 \pm 31$ Mpc. We plot the latter where we use r_s from CMB for plotting purposes.

Our result is in perfect agreement with that obtained by Percival et al. (2009) at $z = 0.275$. Please keep in mind that both results, as well as those obtained by other above studies that analyze the SDSS LRGs, *are not* independent, as we use roughly the same data set. We are encouraged, nevertheless, that our consistency check yields a result in agreement with Percival et al. (2009).

We comment that if we use the median redshift $z_{\rm med} = 0.287$ rather than the mean $\langle z \rangle$ we obtain $r_s/D_V(z = 0.287) = 0.1349 \pm 0.0043$ and $D_V(z_{\rm med}) = 1136 \pm 35$ Mpc.

When comparing these results to cosmological predictions, we assume a flat ΛCDM, and fix $\omega_M \equiv \Omega_{M0}h^2 = 0.1358$. This constraint is motivated by its low (2.7%) uncertainty in the WMAP5 CMB temperature measurements (Komatsu et al. 2009). D_V depends both on Ω_{M0} and h independently, thus its values vary. Meanwhile, r_s depends on ω_M (and $\Omega_b h^2$) so it is kept constant. The 1σ results of r_s/D_V indicate that, constraining w_M from CMB, the preferred region of the matter density lies in the range $\Omega_{M0} = [0.25, 0.34]$, and $h = [0.63, 0.73]$ in agreement with CMB and others. This is consistent for both choices of redshift (mean or median). If we would plot median redshift result ($z = 0.287$) in Figure 11 it would appear along the $\Omega_{M0} = 0.28$ line. We defer a full analysis of the cosmological implications.

Our ξ results for DR7 LRGs as well as the covariance matrix may be be obtained on the World Wide Web.[4]

5. Discussion

The nature of the baryonic acoustic peak and larger scales in $\xi(s)$ have also been discussed in previous studies. Cabré & Gaztañaga (2009) used DR6 (77,000 LRGs) and found a similar level of clustering to ours and examined various possible data analysis systematic effects that might cause the strong signal (the positive ξ at scales larger than $150\,h^{-1}$Mpc). We followed similar steps with the addition that we revisit DR3 to reproduce results from Eisenstein et al. (2005). In agreement with Cabré & Gaztañaga (2009), we show in our Appendix B that data analysis systematics have no significant effects on these results.

The main difference between DR3 and DR7 is the sky coverage. DR7 covers over twice as much sky and, as opposed to the DR3, is continuous. The latter should not be an important issue due

[4]http://cosmo.nyu.edu/∼eak306/BAF.html

to the fact that the random points used in the ξ estimator take into account boundary effects and holes within the RA-Dec plane. The mock catalog tests we conduct suggest that sample variance is a possible explanation for the difference between the large-scale signals of the two data sets.

Martinez et al. (2008) also present a wide baryonic acoustic feature and large-scale clustering in the ξ for SDSS LRGs (DR7). It is worth noting that their definition of M_g is slightly different than the one used here and in Eisenstein et al. (2005). Also, as described in Appendix A.1, we correct for data angular incompleteness where they did not. Nevertheless, our DR7-Full ξ results are in fair agreement with their DR7-LRG.

We show that sample variance affects not only the shape of the signal at large scales (hence helping explain the broadness of the baryonic acoustic feature), but also the probability of detecting a peak: we found that at least 10% of the mock realizations lack a baryonic acoustic signal. Nevertheless, we show, in agreement with other studies mentioned here, that the SDSS LRG sample contains a baryonic acoustic feature which is stable within most redshift and M_g cuts, as well as possible observational bias. Larger surveys are underway to better measure this new holy grail for cosmic distances. For example, the Baryon Oscillation Spectroscopic Survey (BOSS) is estimated to map 1.5 million LRGs in a much larger volume than the DR7, up to $z \sim 0.7$ (Schlegel et al. 2009).

We measure the observed peak position $s_{\rm p}$ to an accuracy of $\sim 3\%$ based on a model constructed from our mock catalogs results. The main source of this uncertainty is due to sample variance, of the DR7-Dim subsample used. Fitting data to a redshift space, non-linear model, we also explain sensitivity of determining the peak position to the range of data points used, as well as shot-noise. These systematics are shown to be less than 2σ of the sample variance, but should be considered when the latter is reduced. Our result $s_{\rm p}$ is in very good agreement with that obtained by Hütsi (2006) who analyzed the wavelength of the Baryonic Acoustic Oscillations in the power spectrum of the SDSS LRGs of DR4.

We use our measurement of $s_{\rm p}$ to determine the ratio r_s/D_V (§4.6). Our result agrees very well with that obtained by Percival et al. (2009) who analyze the oscillations in the power spectrum and quote results at a very similar redshift. Note that we use shape information in the correlation function and do not marginalize over cosmological parameters, but rather tested consistency of one fiducial cosmology. However, our result does not have tighter constraints than that obtained by their study. This is at least partially due to our examination of a smaller volume ($z < 0.36$) to better control systematics (which we test for). It also might be due to our mock catalog uncertainties being larger than the lognormal approximation used in Percival et al. (2009) or a difference in the range of scales used, among other things (see Sánchez et al. 2008 for a comparison of the relative performance of ξ and P(k) estimates).

Regarding claims of the absence of anti-correlations at the largest scales (Labini et al. 2009;

DR7), we point out that the mock realizations show a large variety of crossover values r_c from positive to negative correlations. In Figure 12 we show all 160 DR7-Dim mocks. Their crossover points are indicated by short green lines, and that of the mean (white line) by the orange line at $\sim 140\,h^{-1}$Mpc. We find $\sim 6\%$ (9) of the mocks do not crossover before $200\,h^{-1}$Mpc, but this value should not be taken too seriously as it increases with wider binning, which causes less noisy results and fewer crossovers. A similar result was showed to us by E. Gaztañaga for DR6 mock catalogs. The r_c values (or s_c as we measure in redshift space) are summarized in the histogram in the inset, showing a wide variety of crossovers between $[120, 160]\,h^{-1}$Mpc and even some around $80\,h^{-1}$Mpc. We comment that the crossovers are defined as the first time the ξ crosses through the zero value, and we do not account for returns to positive values. Though having different bias in clustering in respect to matter, galaxies should have the same crossover point between correlation and anti-correlation. We conclude that sample uncertainties still dominate our ability to perform such a test for determining r_c.

6. Summary and Conclusions

Data sets later than DR3 yield a broader baryonic acoustic peak and stronger large-scale clustering signal than measured by Eisenstein et al. (2005). In this paper we have demonstrated that:

1. Differences between DR3 and DR7 results are not due to known systematic uncertainties in data analysis. Applying the same methods in the DR7 analysis, we reproduce the same DR3 data set as Eisenstein et al. (2005) and match the same resulting ξ.

2. Large-scale clustering of DR7 results are in agreement with the ΛCDM+HOD model used here (flat $\Omega_{M0} = 0.25$). The average ξ at scales $130 < s < 400\,h^{-1}$Mpc is within $70 - 95\%$ confidence level according to variation of mock LRG catalogs.

3. The detected baryonic acoustic peak position in DR7 seems stable within most investigated subsamples and agrees with Eisenstein et al. (2005). When analyzing ξ results between $60 < s < 150\,h^{-1}$Mpc we find the peak position to be at $s_{\rm p} = 101.7 \pm 3.0\,h^{-1}$Mpc. This result is sensitive up to 1σ to the upper boundary of range chosen for analysis, random sample shot noise and binning.

4. Using this last result we measure the ratio between the sound-horizon at the end of the baryon drag epoch r_s to D_V (Equation 8) at $r_s/D_V(z = 0.278) = 0.1389 \pm 0.0043$ Mpc. Our result is in excellent agreement with Percival et al. (2009), who analyzes the power spectrum of roughly the same sample, and utilizing a different approach of analysis. Inserting r_s obtained from

WMAP5 we calculate $D_V(0.278) = 1103 \pm 43\, h^{-1}$Mpc. Comparison with other determinations in the literature are also in very good agreement (see Figure 11).

5. We find a lower bound of 10% of mock realizations that do not show evidence of a sign of a baryonic acoustic feature. However, given a mock realization with a clear peak, we conclude that we can measure its peak position value $s_{\rm p}$ within $1\sigma_{s_{\rm p}} \sim 3\, h^{-1}$Mpc.

6. A bright volume-limited subsample of DR7, DR7-Bright, shows a significantly stronger large-scale-signal than predicted by our mock catalogs, which is not explained by potential systematics.

The differences between DR3 and DR7 in the ξ may be explained in two ways: signal variance or observational systematics. Our analysis of the LasDamas mock catalogs show that the signal is still dominated by noise, which yields a variety of large-scale signals, wide and narrow baryonic acoustic peaks, as well as some featureless.

In §4.4 we test a method of how the sensitivity of LRG selection might affect the correlation function. Although Eisenstein et al. (2001) cautions that the number count of LRG targets is sensitive to small variations in color cuts, our test shows robustness of the large-scale clustering, and of the baryonic acoustic feature in particular. Hogg et al. (2005) also demonstrates consistency of the survey calibration in its different patches, by examining DR3 SDSS LRG number counts in different regions of the sky.

Our analysis of the apparent strong large-scale signal of the DR7-Bright subsample compared to the model used does not have the power to rule out clustering predicted by ΛCDM, as it is tested against one cosmology, and one HOD model. As explained, the HOD parameters, were fit using only small-scale clustering, and might not be the best choice on large scales. However, it is reassuring that even in this case, the largest discrepancies of the data compared to our mocks are not much larger than at the 2σ-level (in order to rule out the cosmology used, one would have to marginalize over all possible HOD fits, which we have not done). Furthermore, a large-scale enhancement of the two-point correlation function for a fixed cosmology may be obtained in models of primordial non-Gaussianity with parameter $f_{\rm NL} > 0$, see e.g. Fig. 10 in Desjacques et al. (2009). It will be interesting to follow up these issues with future surveys that will tighten uncertainties.

In summary, we show in this study that the SDSS LRG DR7 sample is consistent with ΛCDM, and the baryonic acoustic feature is stable within its various datasets and most subsamples.

Funding for the SDSS and SDSS-II has been provided by the Alfred P. Sloan Foundation, the Participating Institutions, the National Science Foundation, the U.S. Department of Energy, the National Aeronautics and Space Administration, the Japanese Monbukagakusho, the Max

Planck Society, and the Higher Education Funding Council for England. The SDSS Web Site is http://www.sdss.org/.

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

Fermilab is operated by Fermi Research Alliance, LLC under Contract No. DE-AC02-07CH11359 with the United States Department of Energy.

It is a pleasure to thank Ariel Sánchez for his detailed comments on our manuscript. We thank Daniel Eisenstein for assistance in selecting LRGs and Idit Zehavi for discussions on weighting algorithms and selection function considerations. We also thank Raul Angulo, Enrique Gaztañaga, David Hogg, Michael Joyce, Vincint Martinez, Nikhil Padmanabhan, Will Percival, Ramin Skibba, Francesco Sylos Labini, and Martin White for useful discussions and insight. We thank the LasDamas project for making their mock catalogs publicly available. We thank Horizon team for making their mocks public, and in particular Changbom Park and Juhan Kim for discussions on usage. We thank the referee for his/her comments and suggestions. E.K thanks Adi Nusser for his hospitality at the Technion, Israel Institute of Technology, and Betty K. Rosenbaum for her comments. E.K was partially supported by a Google Research Award. M.B was supported by Spitzer G05-AR-50443, NASA Award NNX09AC85G. R.S. was partially supported by NSF AST-0607747 and NASA NNG06GH21G.

A. Selection Functions

Here we explain the angular and radial selection functions of the SDSS LRG sample, as well as the mock catalogs. When selecting and weighting LRGs in DR3 and DR7, we followed the procedures described in Zehavi et al. (2005) for the most part, and explain here a few differences from that study, none of which affect the end results.

A.1. Angular Mask

The SDSS LRGs cover roughly 20% of the sky, mostly in the Northern Galactic Cap, and partially in the Southern (Figure 1). After an image of the sky is obtained in the u, g, r, i, z bands, objects are targeted for a spectroscopic observation to measure their redshifts. Here we refer to targeted quasars, Main galaxies and LRGs as *targeted objects*. Zehavi et al. (2005), in their appendix, discuss reasons that some targeted objects fail to receive spectra, and their method (which is also used in Eisenstein et al. 2005) of correcting for this small incompleteness. Please refer to their study for definitions of "sector", "sector completeness", "fiber-collisions" and the polygon method with which the complicated angular geometry is expressed in DR3 and later releases.

We choose a slightly different definition for sector completeness than Zehavi et al. (2005). We define it as the number of targeted objects (Quasars+LRG+Main) that have obtained spectra (after fiber-collision corrections) divided by the total number of targeted objects. By "fiber-collision corrections" in the number of targeted objects which obtained spectra, we mean that before calculating sector completeness, we up-weight objects with spectra which are within $55''$ of objects that did not. For example, if we find a group of 4 objects in $55''$ proximity and three of them get spectra, all three are fiber-collision weighted by 4/3. This up-weighting plays two roles in our algorithm: defining sector completeness and while counting galaxy pairs.

We mentioned above our definition for sector completeness. If, for example, the only objects which did not get spectra within a sector are due to fiber-collisions, the sector is considered fully complete. A rare peculiarity occurs due to $55''$ neighbors at straddling the border between sectors, in which one (or more) obtain spectra and another does not. In these cases some sectors might have a completeness larger than unity, meaning they obtain a partial completeness fraction from the neighboring sector. In summary, the average completeness of all sectors is 98% for DR7. If we define Main+LRG as 'targeted objects' (excluding quasars) the completeness yields the same, and if we define targeted objects as LRGs only we obtain 96%. We find that these different definitions for completeness result in subtle mismatches of the DR3 LRGs compared sample to Eisenstein et al. (2005), all of which less than 0.7% of the $\sim 47,000$ galaxies, and negligible effects on ξ compared to shot-noise of random points used. As in Eisenstein et al. (2005) we limit ourselves to sectors

with $> 60\%$ completeness. This results in 29 sectors (that have a non-zero completeness) with a total area of 13 square degrees (0.16% of targeted sky) and a total of 364 targeted LRGs. As in Eisenstein et al. (2005) we used the completeness as a probability to exclude random points from each sector. They also up-weight both data and random points in each sector by the reciprocal of the completeness value. With the high completeness of the survey we do not expect differences in the resulting ξ between the methods.

To account for fiber-collisions while pair-counting, we use a slightly different population than before. LRGs get up-weighted if they are $55''$ neighbors of a targeted LRG (not quasar or Main galaxy) which did not get spectra. This effectively increases our sample LRG number by $\sim 2.2\%$ for DR7-Full and $\sim 1.8\%$ in DR3, which is important to take into account when calculating normalized number of pairs DD, DR and RR.

The LasDamas mock catalogs match the survey geometry as described by the polygon description in the NYU-VAGC. Since the "gamma" release mocks do not model fiber collisions nor missing sectors, completeness is defined to be 100% within the SDSS area. For more details please refer to McBride et. al (2009; in prep).

A.2. Radial Selection Function

The LRG sample used here is the largest quasi-volume-limited spectroscopic sample of its kind today. That said, the radial selection function, $n(z)$ is not constant (meaning not volume limited), as one would expect in a homogeneous universe (up to Poisson noise and radial clustering). Instead it is quasi-volume-limited up to $z \sim 0.36$ and flux limited thereafter. In what follows we show that the features in the $n(z)$ of DR7-Dim do not affect our results. On a more technical note we discuss distributing random points in the radial direction.

We test the LasDamas mock catalogs for the features in the observed radial selection function. The original mocks do not trace the $n(z)$ of the corresponding SDSS redshift and luminosity cuts. The top left inset in Figure 13 shows this difference in shape for the DR7-Dim sample, as LasDamas provides an $n(z)$ with a slight negative slope (cyan; average over 160 realizations). This slope in the LasDamas mocks is a consequence of applying fixed HOD parameters to a light-cone halo catalog, as it neglects the evolution of the dark matter halo mass (i.e. lower halo masses as higher redshift result in fewer artificial galaxies and hence the negative slope). The survey result (thick green histogram) is the same as in Figure 2. Two features noticeable are the negative slope between $0.16 < z < 0.28$, and the positive slope to the peak at $z \sim 0.34$. We show that these features have a negligible effect on ξ and the r.m.s σ_ξ compared to the distribution given by the original mocks, which is closer to volume limited. To validate this claim, we exclude mock LRGs such that

they match the SDSS radial selection function. For the DR7-Dim this meant excluding 15% of the mock galaxies ($n(z)$ average over 160 realization shown as black histogram) and for DR7-Bright it meant excluding 7%, yielding, on average, a similar number count and volume density as in Table 1. In the main plot of Figure 13 we compare results before and after exclusion of the mock galaxies, and find that the mean ξ of 160 realizations agrees on all scales. The right inset shows that the diagonal terms of the covariance matrix $\sigma_\xi \equiv \sqrt{C_{ii}}$ is slightly higher for the galaxy-excluded sample as expected from slightly larger Poisson noise. We conclude that the shape of observed $n(z)$ for the DR7-Dim sample does not appear to affect the results obtained in this study.

We would also like to address the issue of distributing redshifts to random catalogs, a non-trivial step in using random-point based ξ estimators, when dealing with non-volume-limited data. Two popular methods for distributing redshifts (or comoving distances) to random points are redistributing the actual data redshifts randomly, and assigning random distances, so that the overall $n(z)$ shape matches that of the data (i.e, using the data $n(z)$ as a probability function). We test both random point distribution with and without radial weighting.

We show results of both methods of distributing random points in Figure 14. To clarify differences, in the top panel we plot $s \cdot \xi$ and the ratio between all cases to that chosen in this study (weighted Data-$n(z)$; ξ_{ref}) in the bottom panel. For convenience we define the case of distributing data redshifts to random points as "Data-Redshift" (diamonds), and the case of distributing random points to match the data $n(z)$ as "Data-$n(z)$" (crosses). The data used in this analysis are the original DR7-Dim mocks averaged over 8 realizations. We compare results without radial weighting (black) as well as with (bright green, shifted by $2\,h^{-1}$Mpc for visual clarity, shift not calculated within $s \cdot \xi$). The radial weighting when pair counting is explained bellow. The figure shows that "Data-Redshift" is noticeably weaker than "Data-$n(z)$" until $s \lesssim 110\,h^{-1}$Mpc. This is a clear display of the diminishing of the radial clustering modes. Although this effect is very small relative to our current 1σ measurements, it could be important when these are reduced in future surveys. On the larger scales the differences are diminished. The red crosses emphasize the importance of weighting the random points in the same fashion as the data. Not doing so yields spurious clustering.

In conclusion, we choose to radially distribute random points to match the same $n(z)$ shape of the data, using the same weight, and *not* distributing the data redshifts to the random points, as to preserve the radial modes.

After deciding to use the survey $n(z)$ to analyze the mock catalogs we have two options of use of random points: (a) compare all data points to one random set, which has the mean $n(z)$ of the data; or (b) imitate the observation for each realization by comparing each to a random catalog with a tailor-made random catalog. By "tailor-made" we refer to the adjustment of the redshifts z, so that each random catalog yields a similar $n(z)$ to that of the mock data. We applied this

difference when analyzing DR7-Dim and DR7-Bright and find no difference in the ξ mean results. As for the r.m.s σ_ξ we find very small differences in both. On scales up to $s < 200\,h^{-1}$Mpc the DR7-Dim shows fluctuations smaller than 4% where DR7-Bright shows fluctuations below 8%, with no particular preference for either method. This might be a result of the fact that the transverse modes dominate the radial ones. For completeness we note that in our mock results we use the former method, i.e, we apply one random catalog for all the mock data, where the ratio of random to data is $r \sim 50$.

To weight the data approximately according to volume, we need to account for the radial selection function $n(z)$ (see Appendix A.2 and Figure 2). To do so, we apply the standard weighting technique (Feldman et al. 1994).

For the SDSS LRGs we calculate $n(z)$ in bins of $\Delta z = 0.015$ and use a spline interpolation to calculate the radial selection value of each LRG and random point. Each point is then assigned a radial weight of $1/(1 + n(z) \cdot P_w)$ where we use $P_w = 4 \cdot 10^4 h^{-3}\mathrm{Mpc}^3$, as in Eisenstein et al. (2005). The weights are applied to both the data and the random catalogs. The choice for Δz does not affect the measured ξ (as also shown in Cabré & Gaztañaga 2009). We also examined for differences between choosing the observed $n(z)$ for weighting and the model used in Zehavi et al. (2005) and found no difference.

B. Systematics of $\xi(s)$

Here we address possible data analysis effects on our results. We focus on weighting schemes both in the radial and in the angular masks. In general, we find that the position and width of the baryonic acoustic feature does not change much, though we do find small effects in s_p due to Poisson shot-noise.

B.1. Random Shot Noise

When comparing between different ξ results obtained by various systematics, it is important to differentiate biases from variances obtained by random shot-noise.

Applying five different random catalogs to the observed DR7-Dim, at a ratio of ~ 15.6 random points for every data point, and assigning radial weighting (similar to that performed in the analysis) we show in Figure 15 that the random shot-noise is minimal on most scales of concern. The top panel shows ξ for five different random catalogs against the same observed sample (black-dotted, green-dashed, red-dot-short-dashed, orange-long-dashed, blue-triple-dotted-dashed). The bottom panel

shows the ξ ratios with respect to the first random catalog. For the chosen binning, the baryonic acoustic region seems to yield less than a 10% difference. We reran our $s_{\rm p}$ algorithm on all random catalogs and find the peak positions $[101.6, 101.8, 103.4]\,h^{-1}$Mpc. This shows that, although the overall shape seems very consistent when using the different random catalogs, the random-shot noise has a small effect on pinpointing peak position. This is currently smaller than the survey 1σ, but should be considered when statistical uncertainties improve. The fitting normalized χ^2 range for all five catalogs is $\chi^2/d.o.f = [1.24 - 1.58]$. To reduce this effect on our measurement in §4.5 we used a ratio of number of random to data of ~ 50. The peak position for is marked by the top arrow at $s_{\rm p} = 101.7\,h^{-1}$Mpc.

In the region between $50 < s < 90\,h^{-1}$Mpc we see up to 10% differences, which is expected in a sharp logarithmic slope. On smaller scales $s < 50\,h^{-1}$Mpc differences yield up to 3% difference in amplitude.

B.2. Effects of Weighting

In order to optimally measure the correlation function and to account for the fiber collision effects, we apply weighting algorithms. All differences due to choices about how we weight turn out to be much smaller than current 1σ variances. Angular weighting schemes (fiber-collision correction and sector completeness) are explained in Appendix A.1. Here we explain our algorithm for the radial weighting and show results for both.

Figure 16 shows the effects of various angular and radial weighting schemes on the large-scale ξ of DR7-Full (for clarity we present $s \cdot \xi$ in the top panel). The weighting schemes compared are:

1. No weighting at all (black)
2. Radial weighting only (cyan; shifted by $1.75\,h^{-1}$Mpc for clarity)
3. Radial weighting + fiber-collision (red; shifted by $3.5\,h^{-1}$Mpc)

As mentioned in Appendix A.1, sector completeness is taken into account in the distribution of the random points. The bottom panel shows the ratio of each of the above options ξ_i over that chosen in this study (Fiber-Col+Rad; ξ_{ref}). To avoid possible random shot noise (see Appendix §B.1) we use the same set of random points in each case. The weight is modified for each option for the data and random as indicated in the plot. The fiber-collision weight for the random set is always $w = 1$.

We clearly see from the three above options the effect of radial weighting on scales $s \lesssim 170\,h^{-1}$Mpc. We can grossly divide this range into a smaller one of $s \lesssim 115\,h^{-1}$Mpc in which the radial weight

adds some power to the signal, and a larger scale region in which the signal is reduced. The fiber-collision correction weights ("Fiber-Coll") are important when dealing with scales $s \lesssim 40\,h^{-1}$Mpc at the few percent level, but not at the scales discussed in this study. Beyond $170\,h^{-1}$Mpc the effects of the weighting schemes used here are minimal. Most importantly, as shown in Cabré & Gaztañaga (2009), the apparent strong ξ signal at large scales, and the baryonic acoustic feature position are consistent among the various weighting methods.

These weighting effects are not substantial in Figure 14, because the $n(z)$ used there (full DR7-Dim LasDamas catalog without dilution to match SDSS $n(z)$) does not have the same complex features seen in DR7-Full (the mocks used in that figure are instead close to volume-limited).

B.3. Various Systematics

We have checked a few other possible systematics collectively described here.

The SDSS DR7 sample is mostly contiguous, with only 9.8% of the surveyed area not contained in the main part of the Northern Galactic Cap sample (Figure 1). We examined three choices for which footprint to use: the full survey (100% of LRG sample), the Northern Galactic Cap only (90.9%), and the Northern Galactic Cap without the small "island" (90.2%). We found no significant difference in the resulting DR7 ξ for these samples.

In order to calculate the separation in comoving space between galaxy pairs, we must use the observed redshifts and assume a certain fiducial cosmology. For most of this study we assumed a flat ΛCDM model with $\Omega_{M0} = 0.25$. When analyzing differences between DR3 and DR7-Full we also tested the cosmology assumed by Eisenstein et al. (2005) $\Omega_{M0} = 0.3$. We also examine each cosmology on each subsample and found that the cosmology does not affect the resulting ξ or the measured $s_{\rm p}$ significantly relative to our other uncertainties. For a direct comparison of both cosmologies in DR7 see Figure 3.

We also tested how the choice of cosmology affects the selection of LRGs, due to the fact that M_g is sensitive to cosmology. Within DR7-Dim ($0.16 < z < 0.36$, $-23.2 < \mathrm{M_g} < -21.2$) we count $61,899$ LRGs when using $\Omega_{M0} = 0.25$ and $61,102$ when $\Omega_{M0} = 0.30$, a 1.3% difference. When probing the full sample, DR7-Full, we find an agreement in number of LRG selected in the two cosmologies to better than 1%. This implies that for large range of redshift an approximate cosmology should not change the number count very much.

As mentioned in §3, the ξ calculation requires the choice of a particular estimator. We tested the Landy & Szalay (1993) estimator against those proposed by Hamilton (1993), Davis & Peebles (1983), Peebles & Hauser (1974). Figures 17 and 18 show our ξ and σ_ξ results, when using DR7-

Dim. In Figure 17 the mock mean (over 160 realizations) are the lines, while the SDSS result are the symbols. The different colors indicate the different estimators used, where Landy & Szalay (1993) (black solid, LS93 hereon) and Hamilton (1993)(green triple-dot-dashed, HAM93) are not distinguishable by eye for the most part. The inset shows the noise-to-signal ratio. Bear in mind that the spikes around $140\,h^{-1}$Mpc simply result from the signal crossing zero around that scale. In Figure 18 the notation is the same when plotting $\sigma_\xi(s)$ for the different estimators. The inset in this plot shows the ratio of each estimator relative to that of LS93:

We find the following:

1. Mean value $\overline{\xi}$: Averaging over all 160 mocks we see that LS93 and HAM93 agree on all scales. The inset of Figure 18 shows that Peebles & Hauser (1974) (PH74) deviates from the latter by 10% at scales of $\sim 75\,h^{-1}$Mpc and Davis & Peebles (1983) (DP83) does the same at $85\,h^{-1}$Mpc. We also note that LS93 and HAM93 asymptote faster to zero than the others.

2. r.m.s σ_ξ: The r.m.s of the signal varies among the estimators on various scales. PH74 yields the same σ_ξ as LS93 up to scales of $\sim 40\,h^{-1}$Mpc before it starts strongly branching off. This is clearly seen in the main plot of Figure 18, and indicated in the other plots by the strong variation of the observed result. Both HAM93 and DP83 have a larger variance than LS93 at scales smaller than $10\,h^{-1}$Mpc. HAM93 later matches the LS93 on all larger scales very well, where DP83 breaks off at $\sim 60\,h^{-1}$Mpc.

We conclude that the Landy & Szalay (1993) and Hamilton (1993) estimators agree in signal for each realization on large scales, and perform much better than the other two as their variance is smaller, and converge much more quickly to zero. In particular, the other two estimators yield very large uncertainties on large scales for individual realizations. We also find that Hamilton (1993) does not perform as well as Landy & Szalay (1993) on smaller scales $s < 10\,h^{-1}$Mpc. Our analysis of DR7-Dim agrees with the other samples DR7-Bright, DR7-Bright-Near, and DR7-Bright-Far. Here we use ratio of random points to data of $r \sim 15.6$ for the SDSS and $r \sim 10$ for mocks. Kerscher et al. (2000) shows Landy & Szalay (1993) to be the estimator with the best performance relative to the true ξ in periodic box measurements, and Labini et al. (2009) shows similar comparison results to ours.

To test the importance of binning, we use the Landy & Szalay (1993) on all 160 mock DR7-Dim realizations and find a strong dependence of the variance on the bin size, as expected. We test various bin widths Δs between $\Delta s = [0.55, 10]\,h^{-1}$Mpc noting that variance changes strongly, where the lowest Δs yields the noisiest σ_ξ, due to the higher shot-noise in each bin. For most of our analysis, including the baryonic acoustic feature, we used $\Delta s = 4.44\,h^{-1}$Mpc. We find that peak position $s_{\rm p}$ is not altered by more than 1σ for any choice of bin size.

We also check for boundary limits in our survey, to get an idea of at what point our ξ estimates are encountering the edges of the survey. We show in Figure 18 a negative slope for σ_ξ all the way to $s < 400\,h^{-1}$Mpc. A region we definitely want to avoid in analysis is one in which σ_ξ has a positive slope, which indicates an upper limit to the effective scaling given the survey volume. For this reason in the top panel of Figure 19 we continue this plot for DR7-Dim (black solid) and DR7-Bright-Near (green triple-dotted-dashed) to $800\,h^{-1}$Mpc . These subsamples (both limited to the range $0.16 < z < 0.36$) have a declining σ_ξ to $500\,h^{-1}$Mpc, and a slightly positive slope thereafter. In the second plot from the top we show the normalized data-data pairs DD of DR7-Dim and DR7-Bright-Near (same notation), as well as the random-random pairs RR count of DR7-Dim (blue dot-dashed). The dip feature at $\sim 500\,h^{-1}$Mpc mentioned before appears here as the peak at that scale. In the third plot from top we differentiate the previous results by s to better see where the number of pairs stop growing with radius. We see a clear crossover between $500 - 600\,h^{-1}$Mpc.

The bottom two plots of Figure 19 can be considered "sanity checks", as we verify basic statistics. Given a periodic box we expect the number of random points N around a given a point at radius s in a spherical shell of width ds to be: $N = 4\pi s^2 ds\overline{n}$ where $\overline{n}$ is the mean density. $RR(s)$, is expected to go as $\sim N(s) \cdot N_R/2$, where N_R is the number of random points, and 2 takes into account double counting. Hence $d\ln(RR)/d\ln(s)$ should yield 2. Deviations from this result are due to boundary effects. In the fourth from top plot we present $d\ln(x)/d\ln(s)$, where x is RR (blue dot-dashed) and DD (solid black). The important result from this plot is the deviation of $d\ln(RR)/d\ln(s)$ from 2 at large-scales. A 5% difference is noticed at $50\,h^{-1}$Mpc. At the baryonic acoustic feature scale the deviation from 2 is $\sim 10\%$, at $200\,h^{-1}$Mpc by $\sim 23\%$ and at $400\,h^{-1}$Mpc by 60%. Although this panel indicates our volume is far from ideal for analysis of $400\,h^{-1}$Mpc, the LS93 estimator appears to be valid as it corrects for these boundary effects by comparing number of data pairs to number of *expected* random points given boundary conditions. We performed the same test for random points within a volume of DR7-Bright (red dot-dashed) where we extend the measurement to $500\,h^{-1}$Mpc.

In the bottom plot we determine how many cubes of length s would fit into the volume contained within DR7-Dim (black solid) and DR7-Bright (red three-dot-dashed). Our choice of $s = 400\,h^{-1}$Mpc yields 10.31 for the former and 18.65 for the latter. For DR7-Bright up to $s = 500\,h^{-1}$Mpc we count 9.55 cubes. As for the peak position at $s = 101\,h^{-1}$Mpc, we count 641 cubes for the DR7-Dim volume and 1159 for a DR7-Bright volume.

All the plots in Figure 19 indicate that our choice of $s < 400\,h^{-1}$Mpc is a fair one within the given sample, and does not depend too much on edge effects.

We also test for uncertainties σ_ξ of the observed DR7-Dim and DR7-Bright against those predicted by Gaussianity plus shot noise (Bernstein 1994) at large scales,

$$\sigma_\xi^2(s) = \frac{2}{(2\pi)^3 V} \int d^3k \; [P(k) + \bar{n}^{-1}]^2 \, j_0(ks)^2, \tag{B1}$$

where V is the volume of the survey (see Table 1), $P(k)$ the galaxy power spectrum, and j_0 the spherical Bessel function of zeroth order. We first test this formula against the variance derived from the mock catalogs for both Dim and Bright samples, and find that at $50 < s < 100\,h^{-1}$Mpc the variance is consistent with Eq. B1. At smaller scales non-Gaussian contributions to the uncertainties quickly make Eq. (B1) an underestimate of the uncertainties, whereas at scales $s > 100\,h^{-1}$Mpc contributions to Eq. (B1) from edge effects (not included there) become important, e.g. making the total uncertainty about 20% larger than given by Eq. (B1) at $300\,h^{-1}$Mpc. Since edge effects depend only on the geometry and density of the sample, we can extract their value from this comparison and apply it to the data, which has the same characteristics (but different two-point function).

From the measured two-point function in the data we use Eq. (B1) including the edge effects variance to compute Gaussian uncertainties for a model that matches the observed ξ. We find uncertainties that are very close to those in the mock catalogs at $s > 150\,h^{-1}$Mpc, the main reason being that discreteness contributions are significant. This suggests that had we changed cosmology and HOD to give a better fit to ξ at largest scales, we would have gotten very similar uncertainties, and thus similar $\sim 2\sigma$ deviations away from zero-crossing for the Bright sample.

REFERENCES

Abazajian, K. & Sloan Digital Sky Survey, f. t. 2009, ApJS, 182, 543

Abazajian, K. et al. 2005, AJ, 129, 1755

Adelman-McCarthy, J. K. & Sloan Digital Sky Survey, f. t. 2008, ApJS, 175, 297

Adelman-McCarthy, J. K. et al. 2006, ApJS, 162, 38

Angulo, R. E., Baugh, C. M., Frenk, C. S., & Lacey, C. G. 2008, MNRAS, 383, 755

Bashinsky, S. & Seljak, U. 2004, Phys. Rev. D, 69, 083002

Berlind, A. A. & Weinberg, D. H. 2002, ApJ, 575, 587

Bernstein, G. M. 1994, ApJ, 424, 569

Blake, C., Collister, A., Bridle, S., & Lahav, O. 2007, MNRAS, 374, 1527

Blake, C. & Glazebrook, K. 2003, ApJ, 594, 665

Blanton, M. R., Lin, H., Lupton, R. H., Maley, F. M., Young, N., Zehavi, I., & Loveday, J. 2003, AJ, 125, 2276

Cabré, A. & Gaztañaga, E. 2009, MNRAS, 393, 1183

Cole, S., Percival, W. J., Peacock, J. A., Norberg, P., Baugh, C. M., Frenk, C. S., Baldry, I., Bland-Hawthorn, J., Bridges, T., Cannon, R., Colless, M., Collins, C., Couch, W., Cross, N. J. G., Dalton, G., Eke, V. R., De Propris, R., Driver, S. P., Efstathiou, G., Ellis, R. S., Glazebrook, K., Jackson, C., Jenkins, A., Lahav, O., Lewis, I., Lumsden, S., Maddox, S., Madgwick, D., Peterson, B. A., Sutherland, W., & Taylor, K. 2005, MNRAS, 362, 505

Coles, P. 1993, MNRAS, 262, 1065

Colless, M., Peterson, B. A., Jackson, C., Peacock, J. A., Cole, S., Norberg, P., Baldry, I. K., Baugh, C. M., Bland-Hawthorn, J., Bridges, T., Cannon, R., Collins, C., Couch, W., Cross, N., Dalton, G., De Propris, R., Driver, S. P., Efstathiou, G., Ellis, R. S., Frenk, C. S., Glazebrook, K., Lahav, O., Lewis, I., Lumsden, S., Maddox, S., Madgwick, D., Sutherland, W., & Taylor, K. 2003, ArXiv Astrophysics e-prints

Crocce, M. & Scoccimarro, R. 2008, Phys. Rev. D, 77, 023533

Davis, M. & Peebles, P. J. E. 1983, ApJ, 267, 465

Desjacques, V., Seljak, U., & Iliev, I. T. 2009, MNRAS, 631

Eisenstein, D. J., Blanton, M., Zehavi, I., Bahcall, N., Brinkmann, J., Loveday, J., Meiksin, A., & Schneider, D. 2005, ApJ, 619, 178

Eisenstein, D. J. & Hu, W. 1998, ApJ, 496, 605

Eisenstein, D. J., Hu, W., & Tegmark, M. 1999, ApJ, 518, 2

Eisenstein, D. J., Seo, H.-J., Sirko, E., & Spergel, D. N. 2007a, ApJ, 664, 675

Eisenstein, D. J., Seo, H.-J., & White, M. 2007b, ApJ, 664, 660

Eisenstein, D. J. et al. 2001, AJ, 122, 2267

Eisenstein, D. J. et al. 2005, ApJ, 633, 560

Feldman, H. A., Kaiser, N., & Peacock, J. A. 1994, ApJ, 426, 23

Fry, J. N. & Gaztanaga, E. 1993, ApJ, 413, 447

Fukugita, M., Ichikawa, T., Gunn, J. E., Doi, M., Shimasaku, K., & Schneider, D. P. 1996, AJ, 111, 1748

Gabrielli, A., Joyce, M., & Sylos Labini, F. 2002, Phys. Rev. D, 65, 083523

Gaztañaga, E., Cabré, A., & Hui, L. 2009, MNRAS, 399, 1663

Gunn, J. E., Carr, M. A., Rockosi, C. M., Sekiguchi, M., et al. 1998, AJ, 116, 3040

Gunn, J. E. et al. 2006, AJ, 131, 2332

Hamilton, A. J. S. 1993, ApJ, 417, 19

Hogg, D. W. 1999, astro-ph/9905116

Hogg, D. W., Eisenstein, D. J., Blanton, M. R., Bahcall, N. A., Brinkmann, J., Gunn, J. E., & Schneider, D. P. 2005, ApJ, 624, 54

Hogg, D. W., Finkbeiner, D. P., Schlegel, D. J., & Gunn, J. E. 2001, AJ, 122, 2129

Hütsi, G. 2006, A&A, 449, 891

Ivezić, Ž. et al. 2004, Astronomische Nachrichten, 325, 583

Kerscher, M., Szapudi, I., & Szalay, A. S. 2000, ApJ, 535, L13

Kim, J., Park, C., Gott, J. R., & Dubinski, J. 2009, ApJ, 701, 1547

Komatsu, E., Dunkley, J., Nolta, M. R., Bennett, C. L., Gold, B., Hinshaw, G., Jarosik, N., Larson, D., Limon, M., Page, L., Spergel, D. N., Halpern, M., Hill, R. S., Kogut, A., Meyer, S. S., Tucker, G. S., Weiland, J. L., Wollack, E., & Wright, E. L. 2009, ApJS, 180, 330

Labini, F. S., Vasilyev, N. L., & Baryshev. 2009, ArXiv e-prints

Landy, S. D. & Szalay, A. S. 1993, ApJ, 412, 64

Lupton, R. H., Gunn, J. E., Ivezić, Z., Knapp, G. R., Kent, S., & Yasuda, N. 2001, in ASP Conf. Ser. 238: Astronomical Data Analysis Software and Systems X, Vol. 10, 269

Martinez, V. J., Arnalte-Mur, P., Saar, E., de la Cruz, P., Pons-Borderia, M. J., Paredes, S., Fernandez-Soto, A., & Tempel, E. 2008, ArXiv e-prints

Matsubara, T. 2004, ApJ, 615, 573

Narayanan, V. K., Spergel, D. N., Davé, R., & Ma, C.-P. 2000, ApJ, 543, L103

Okumura, T., Matsubara, T., Eisenstein, D. J., Kayo, I., Hikage, C., Szalay, A. S., & Schneider, D. P. 2008, ApJ, 676, 889

Padmanabhan, N., Schlegel, D. J., Finkbeiner, D. P., Barentine, J. C., Blanton, M. R., Brewington, H. J., Gunn, J. E., Harvanek, M., Hogg, D. W., Ivezić, Ž., Johnston, D., Kent, S. M., Kleinman, S. J., Knapp, G. R., Krzesinski, J., Long, D., Neilsen, Jr., E. H., Nitta, A., Loomis, C., Lupton, R. H., Roweis, S., Snedden, S. A., Strauss, M. A., & Tucker, D. L. 2008, ApJ, 674, 1217

Padmanabhan, N. & White, M. 2008, Phys. Rev. D, 77, 123540

Padmanabhan, N. et al. 2007, MNRAS, 378, 852

Peebles, P. J. E. 1980, The large-scale structure of the universe (Princeton University Press)

Peebles, P. J. E. & Hauser, M. G. 1974, ApJS, 28, 19

Percival, W. J., Cole, S., Eisenstein, D. J., Nichol, R. C., Peacock, J. A., Pope, A. C., & Szalay, A. S. 2007, MNRAS, 381, 1053

Percival, W. J., Reid, B. A., Eisenstein, D. J., Bahcall, N. A., Budavari, T., Fukugita, M., Gunn, J. E., Ivezic, Z., Knapp, G. R., Kron, R. G., Loveday, J., Lupton, R. H., McKay, T. A., Meiksin, A., Nichol, R. C., Pope, A. C., Schlegel, D. J., Schneider, D. P., Spergel, D. N., Stoughton, C., Strauss, M. A., Szalay, A. S., Tegmark, M., Weinberg, D. H., York, D. G., & Zehavi, I. 2009, ArXiv e-prints

Pier, J. R., Munn, J. A., Hindsley, R. B., Hennessy, G. S., Kent, S. M., Lupton, R. H., & Ivezić, Ž. 2003, AJ, 125, 1559

Reid, B. A., Percival, W. J., Eisenstein, D. J., Verde, L., Spergel, D. N., Skibba, R. A., Bahcall, N. A., Budavari, T., Fukugita, M., Gott, J. R., Gunn, J. E., Ivezic, Z., Knapp, G. R., Kron, R. G., Lupton, R. H., McKay, T. A., Meiksin, A., Nichol, R. C., Pope, A. C., Schlegel, D. J., Schneider, D. P., Strauss, M. A., Stoughton, C., Szalay, A. S., Tegmark, M., Weinberg, D. H., York, D. G., & Zehavi, I. 2009, ArXiv e-prints

Richards, G. et al. 2002, AJ, 123, 2945

Sánchez, A. G., Baugh, C. M., & Angulo, R. 2008, MNRAS, 390, 1470

Sanchez, A. G., Crocce, M., Cabre, A., Baugh, C. M., & Gaztanaga, E. 2009, ArXiv e-prints

Scherrer, R. J. & Weinberg, D. H. 1998, ApJ, 504, 607

Schlegel, D., White, M., & Eisenstein, D. 2009, ArXiv e-prints

Seo, H.-J. & Eisenstein, D. J. 2007, ApJ, 665, 14

Seo, H.-J., Siegel, E. R., Eisenstein, D. J., & White, M. 2008, ApJ, 686, 13

Smith, J. A., Tucker, D. L., et al. 2002, AJ, 123, 2121

Smith, R. E., Scoccimarro, R., & Sheth, R. K. 2008, Phys. Rev. D, 77, 043525

Spergel, D. N. et al. 2003, ApJS, 148, 175

Stoughton, C. et al. 2002, AJ, 123, 485

Strauss, M. A. et al. 2002, AJ, 124, 1810

Tegmark, M. et al. 2006, Phys. Rev. D, 74, 123507

Tucker, D. L. et al. 2006, Astronomische Nachrichten, 327, 821

York, D. et al. 2000, AJ, 120, 1579

Zehavi, I., Eisenstein, D. J., Nichol, R. C., Blanton, M. R., Hogg, D. W., Brinkmann, J., Loveday, J., Meiksin, A., Schneider, D. P., & Tegmark, M. 2005, ApJ, 621, 22

This preprint was prepared with the AAS LaTeX macros v5.0.

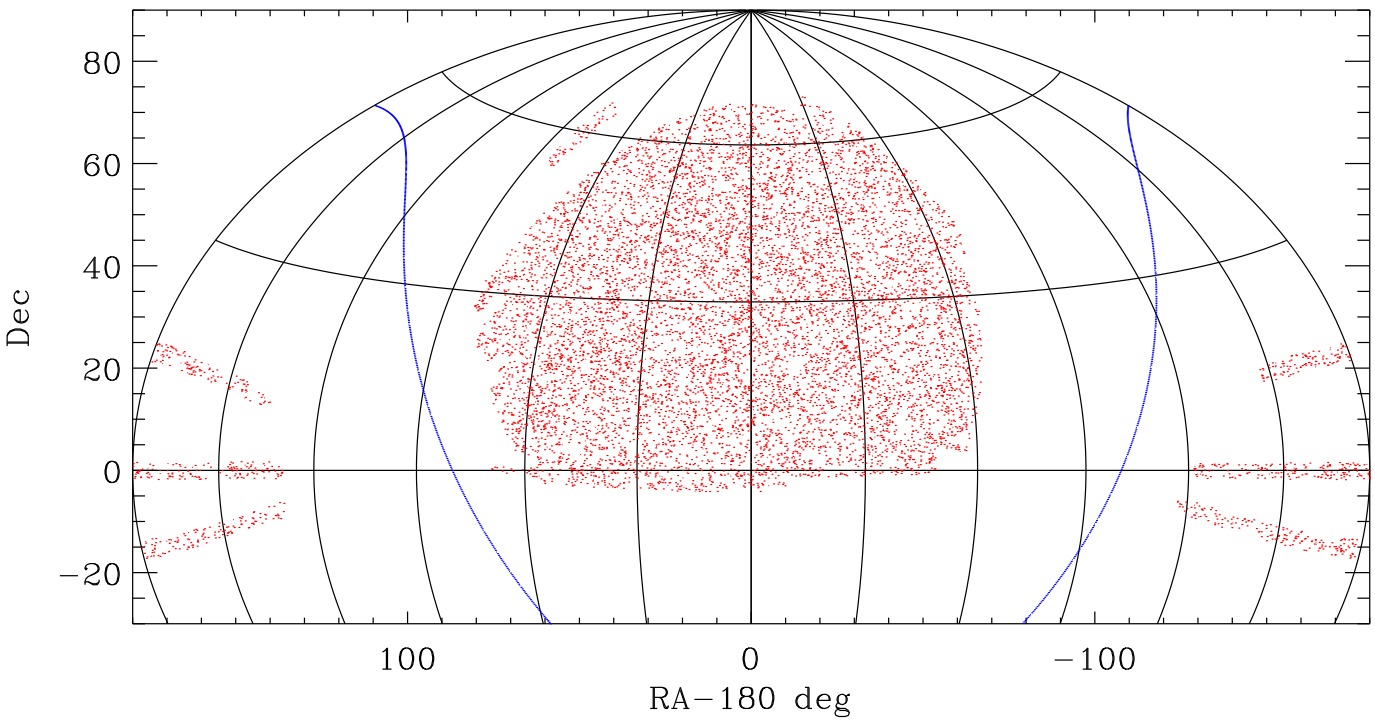

Fig. 1.— Angular Selection Function : SDSS DR7 LRG sky coverage. For plotting purposes we present one tenth of the $105,831$ DR7-Full galaxies. The solid blue line is the Galactic plane.

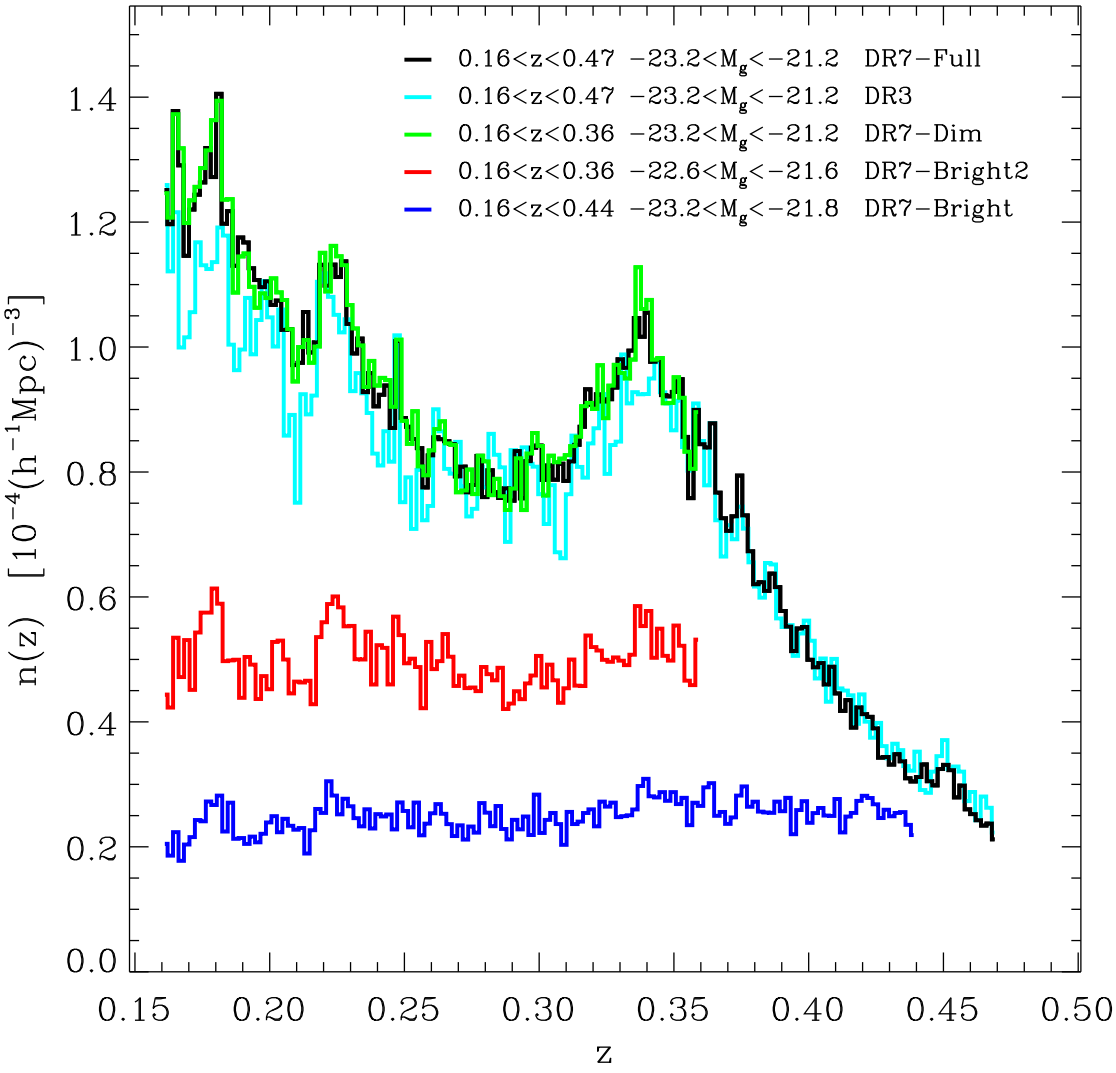

Fig. 2.— Radial selection function: Comoving number density $n(z)$ of the full DR7 (DR7-Full; black) and its subsamples DR7-Dim (green), DR7-Bright (blue), DR7-Bright2 (red) and DR3 (cyan).

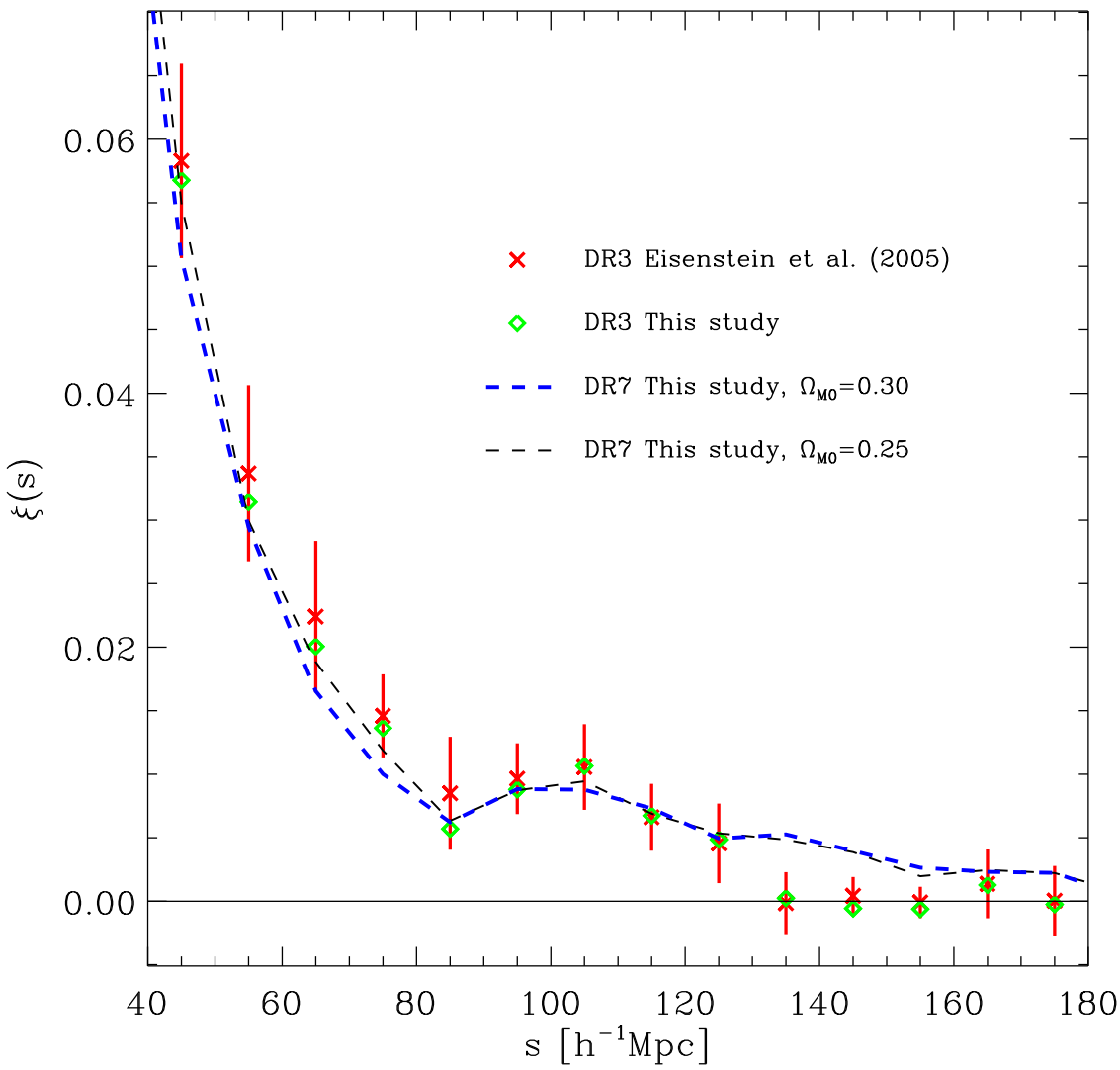

Fig. 3.— DR3 and DR7 $\xi(s)$: Our DR3 results (green diamonds) show excellent agreement on the scales investigated here with those published by Eisenstein et al. (2005) (red crosses and uncertainty bars). The remaining discrepancies are consistent with shot-noise in the random catalogs. The dashed lines are our results for DR7-Full which shows a stronger clustering signal at $135 < s < 180\,h^{-1}$Mpc . In the thick blue dashed line we used the same $\Omega_{M0} = 0.3$ flat cosmology as the DR3 results, and thin black dashed line $\Omega_{M0} = 0.25$. Both cosmologies agree very well at scales discussed here $80 < s < 200\,h^{-1}$Mpc. In our DR3 result we use number ratio of random to data points $r \sim 50$ and for DR7 ~ 15.6.

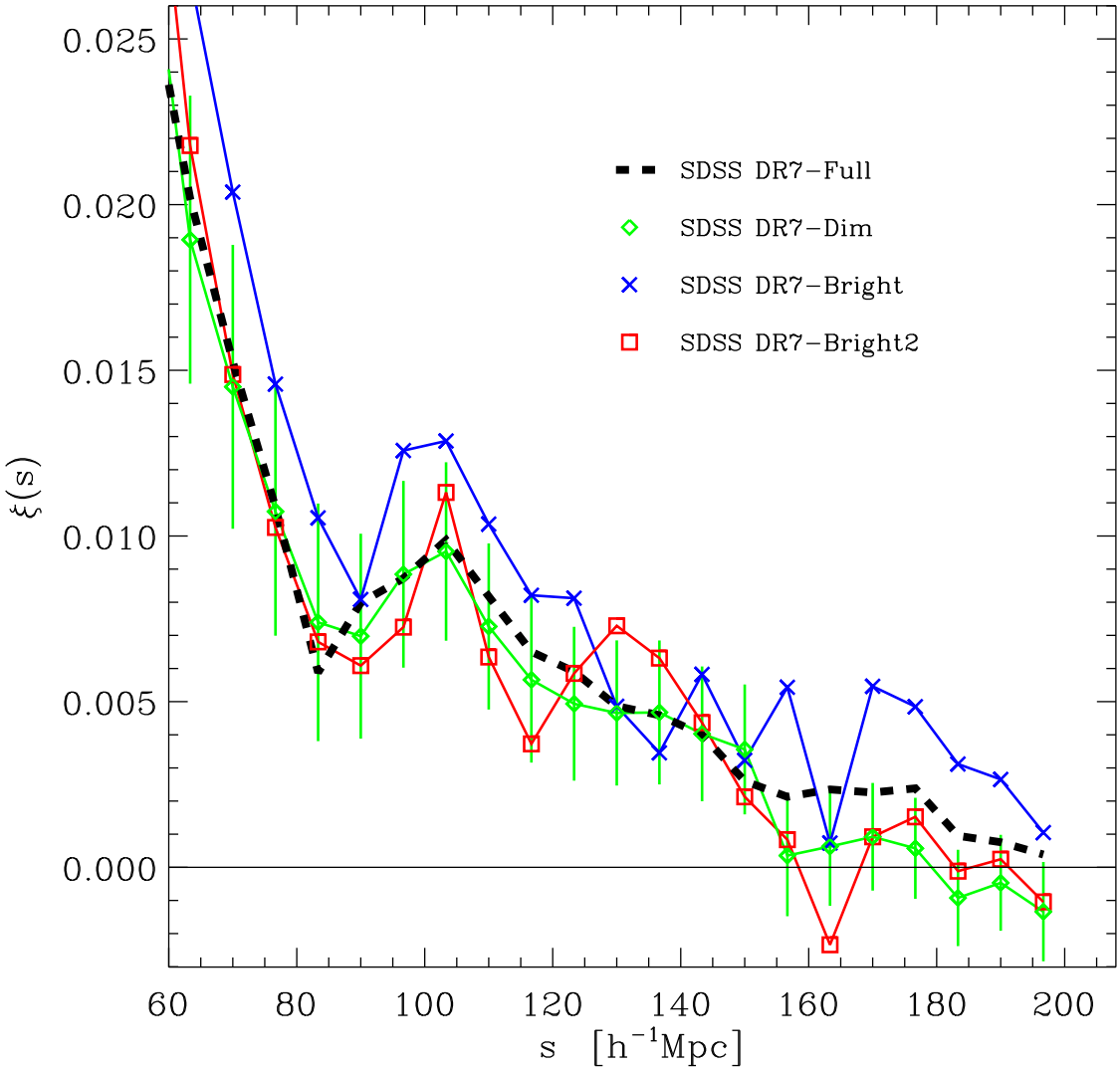

Fig. 4.— DR7 Subsample $\xi(s)$: Comparing DR7-Full (thick dashed line) to subsamples DR7-Dim (green diamonds with uncertainty bars), DR7-Bright (blue crosses) and DR7-Bright2 (red squares). The peak position appears consistent for all subsamples. For the uncertainties of DR7-Dim and DR7-Bright please refer to Figures 5 and 8, respectively.

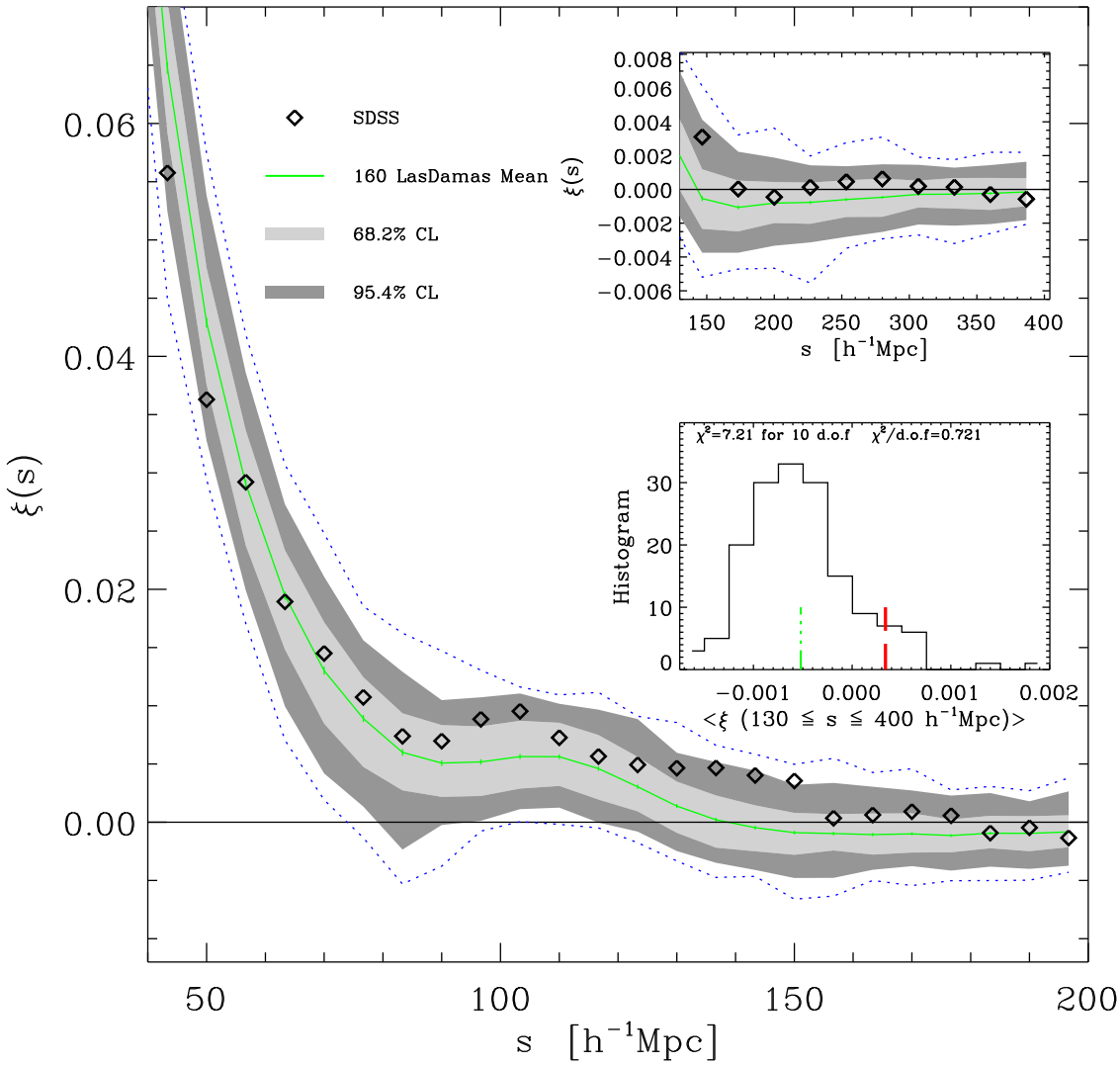

Fig. 5.— DR7-Dim $\xi(s)$: Results from SDSS (black diamonds) and from the LasDamas mock catalogs. The mock mean $\overline{\xi}_{mock}$ is the green solid line and the uncertainties in the mean are small vertical green lines. The variance for one realization is presented by the gray bands: 68.2% light gray ($1\sigma_{mock}$), and 95.4% by dark ($2\sigma_{mock}$). The blue dotted lines are the outermost result of all mocks in each separation bin (not one realization in particular). Top Inset: Same format as main figure extending the results with wider bins to larger scales. Bottom Inset: Significance of large-scale clustering - we average ξ in bins $[130, 400]\,h^{-1}$Mpc . The observed $\langle\xi\rangle$ (red thick dashed line) is clearly within 2σ of the mock realizations (black histogram), with a χ^2 fitting on 10 d.o.f yielding $\chi^2/d.o.f = 0.72$. The mock mean result is the thin green dot-dashed line.

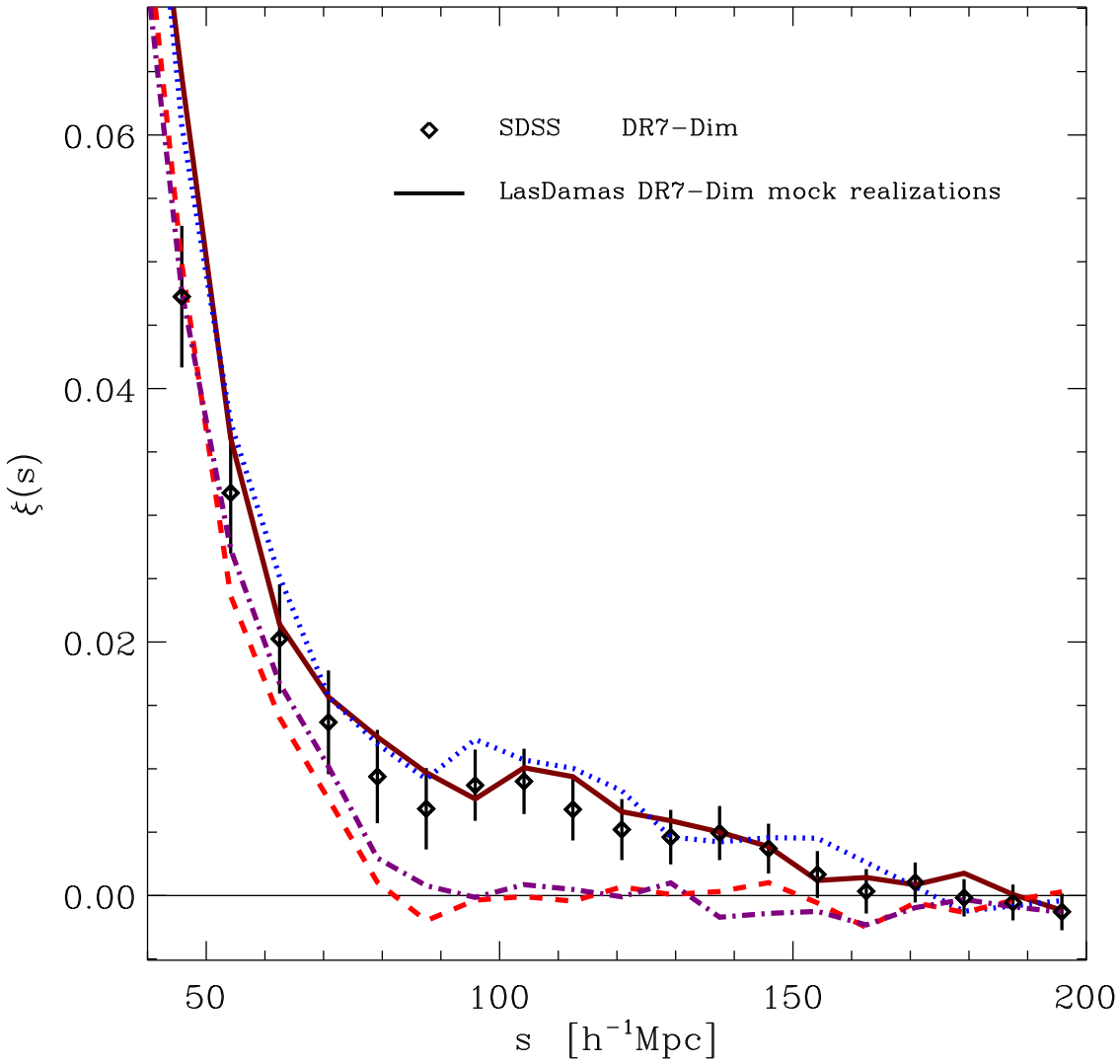

Fig. 6.— DR7-Dim $\xi(s)$ Mock Realizations: Hand-picked LasDamas mocks (lines) compared with the SDSS result (diamonds). The green (dotted) and red (solid) show fairly good agreement on most large-scales with observation, though their peak position appear to be in different locations. The orange (dashed) and cyan (dot-dashed) mocks show example realizations without a baryonic acoustic feature. Using a liberal approach, we counted a minimum bound of 10% (17) from the full set of 160 mocks with no sign of a peak.

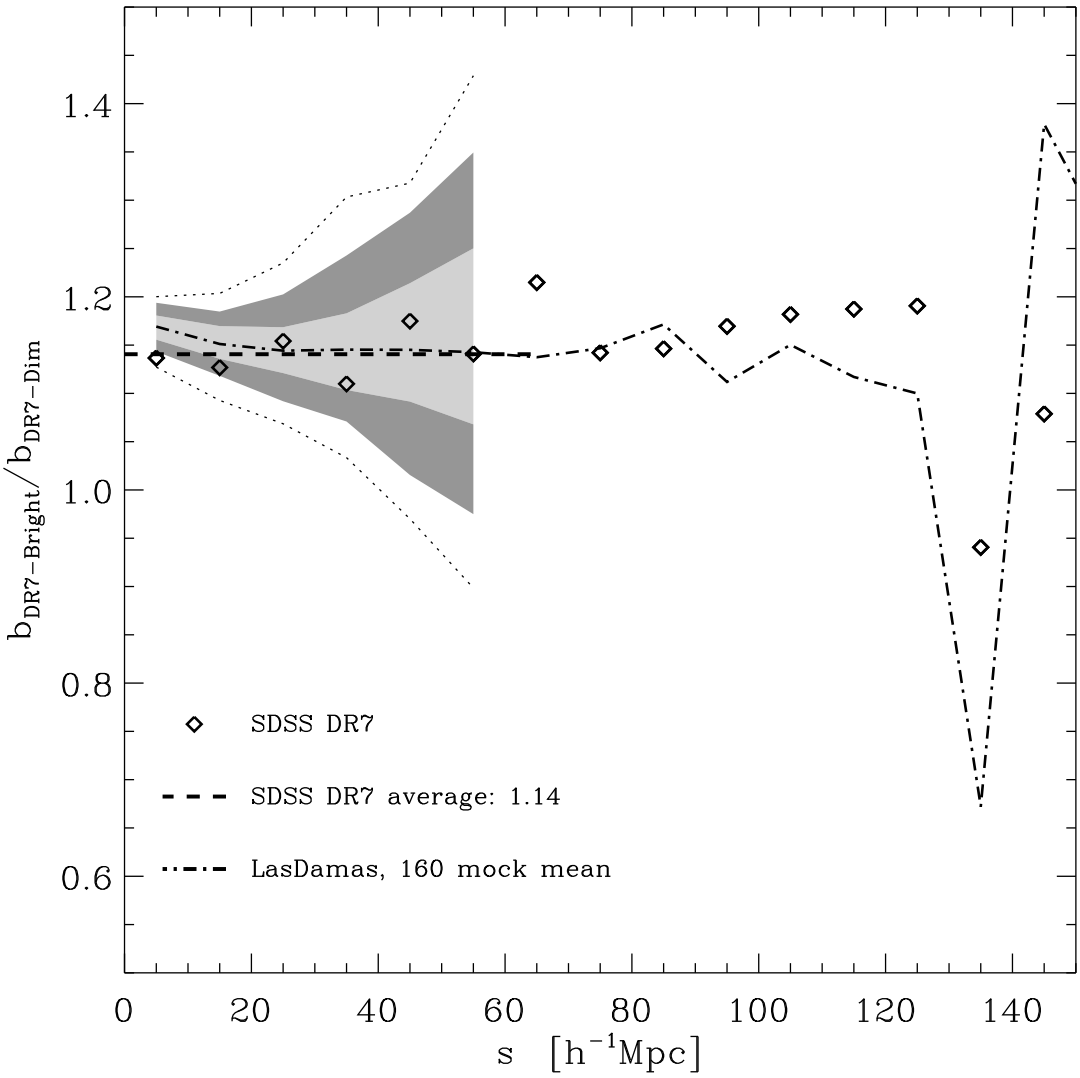

Fig. 7.— Relative Bias: Comparing redshift-space bias ratio $b_{DR7-Bright}/b_{DR7-Dim} \equiv \sqrt{\xi_{DR7-Bright}/\xi_{DR7-Dim}}$ between the two subsamples for both the observed LRGs (symbols) as well as the LasDamas mock mean (dot-dashed lines). The average observed value is taken between $0-60\,h^{-1}$Mpc indicated in thick dashed line valued at 1.14. The gray bands indicate the one and two σ distribution of the relative bias between each DR7-Dim sample and its same realization DR7-Bright sample. The dotted lines are the outer values for each bin. Mock uncertainty calculations indicate that this bias estimate is significant (above 2σ) at $s < 40\,h^{-1}$Mpc.

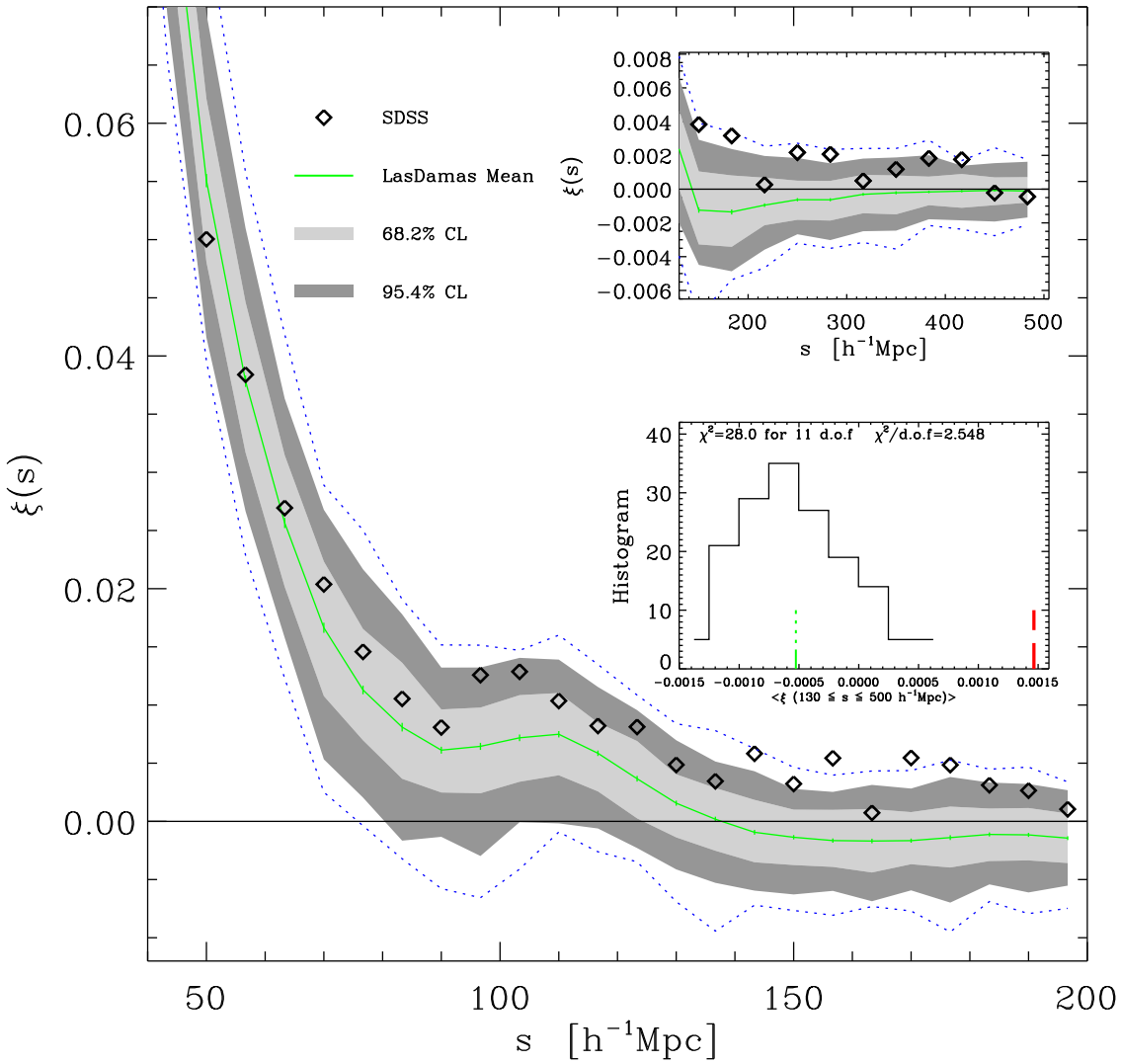

Fig. 8.— DR7-Bright $\xi(s)$: Same format as Figure (5) for the brighter sample. The insets here extend to $500\,h^{-1}$Mpc. The observed subsample shows a significantly stronger signal at large scales than produced by our ΛCDM+HOD model.

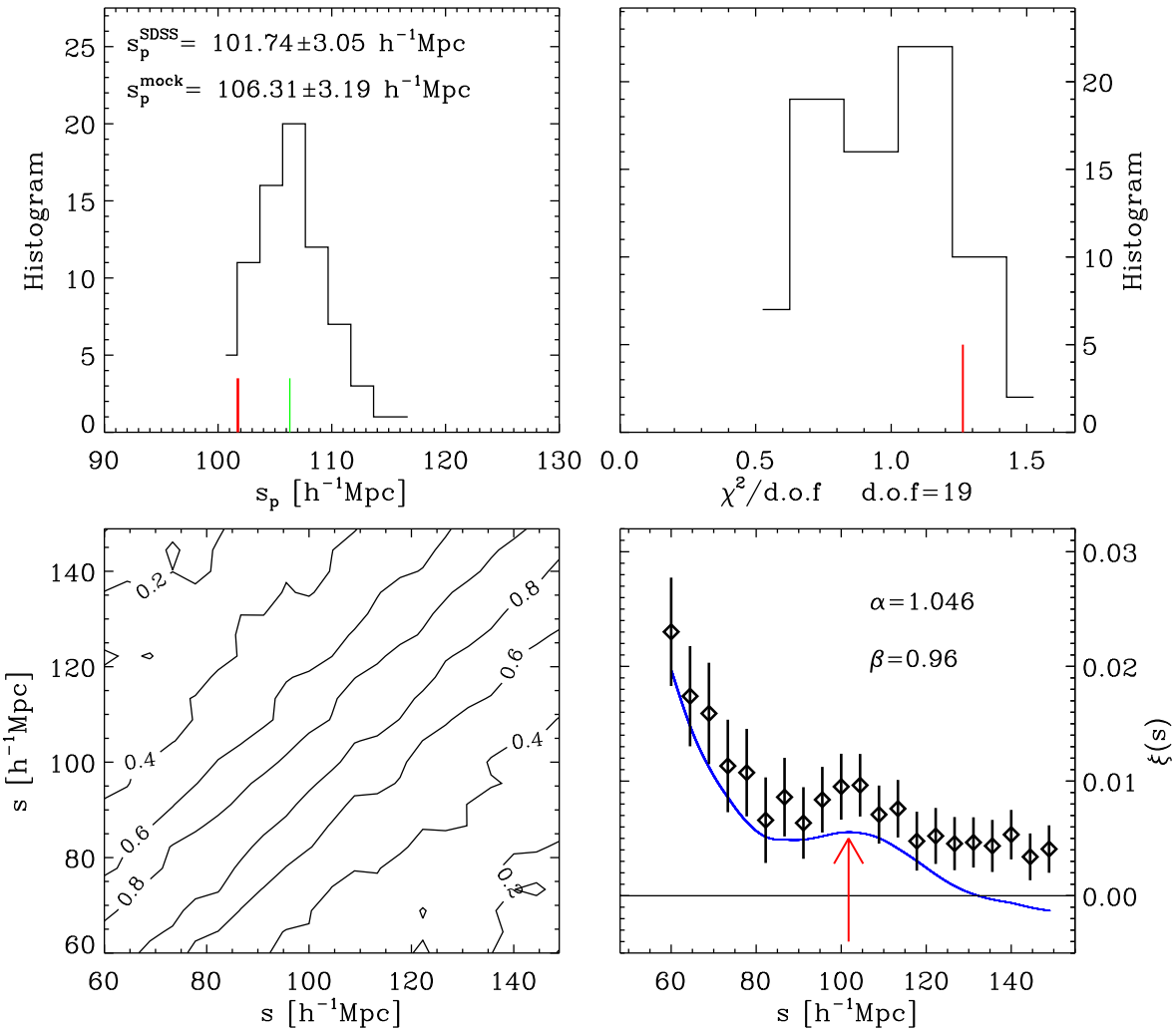

Fig. 9.— Measuring the Baryonic Acoustic Peak Position $s_{\rm p}$ in DR7-Dim Sample: Top Left Panel- $s_{\rm p}$ histogram of 75 mock realizations with clear sign of a peak. Vertical red (thick dark) solid line is observed value and green (thin bright) is that of mock mean. $\sigma = 3.05\,h^{-1}$Mpc is the 68.2% CL around the mock peak $s_p^{mock} = 106.31\,h^{-1}$Mpc normalized to the observation by $\sigma \sim s_p$. Top Right Panel- χ^2 histogram for mock realization fits to model, where the red line indicates that of the observation at $\chi^2 = 23.9$ for 19 d.o.f. Bottom Right Panel- Fitting the observed $\xi(s)$ (black diamonds) to model $\beta\overline{\xi}(\alpha \cdot s)$ (blue line), where $\overline{\xi}$ is the mock mean. Red arrow indicates $s_{\rm p}$= 101.7 h^{-1}Mpc and is the same as in the top left panel. Indicated also are best-fit α and β parameters. Black vertical lines are $\sqrt{C_{ii}}$. Bottom Left Panel- $C_{ij}/(\sigma_i\sigma_j)$.

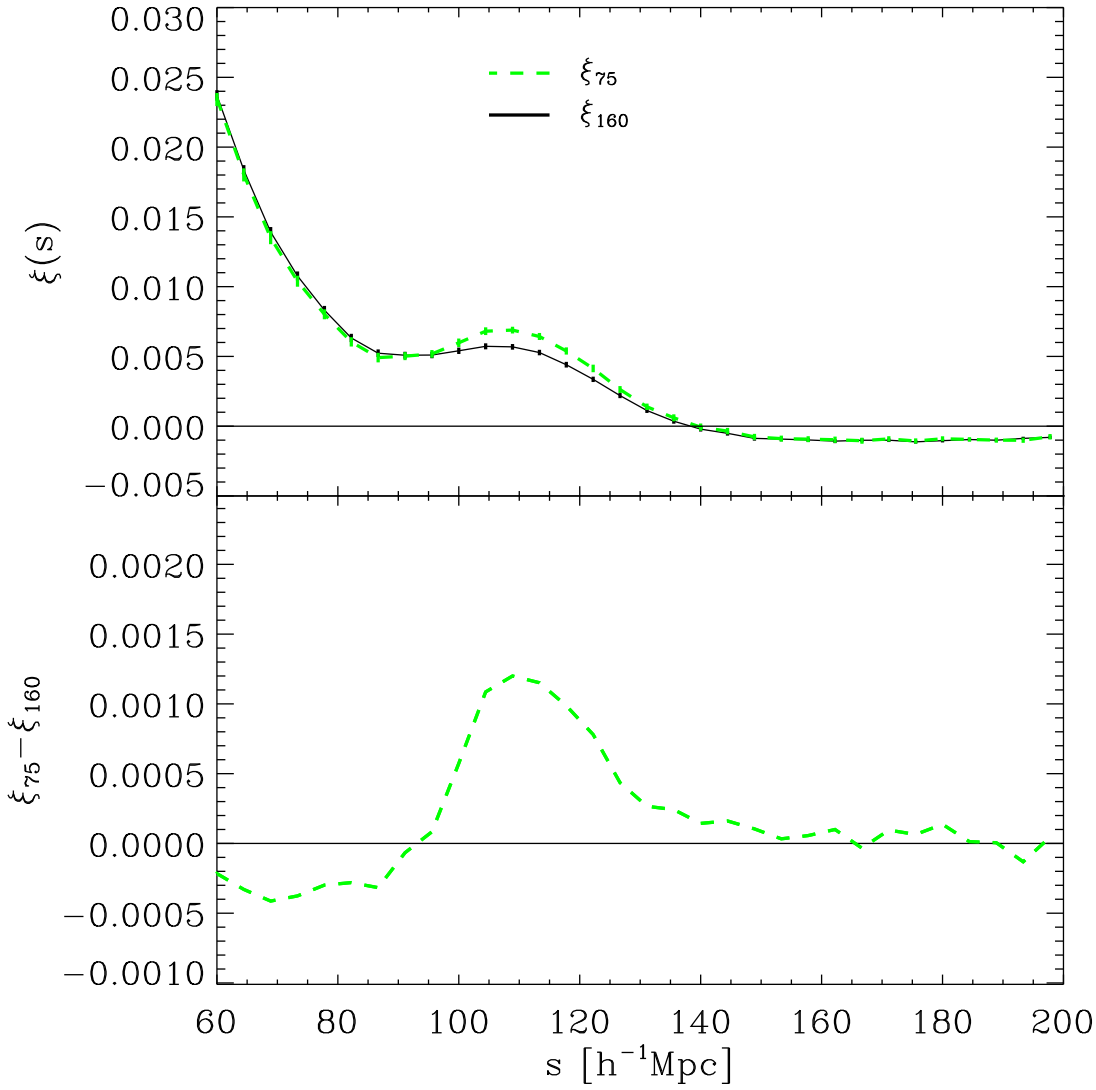

Fig. 10.— Baryonic Acoustic Peak Position in DR7-Dim mocks: Top Panel- The line for the 75 clear-peaked mocks $\overline{\xi}_{75}$ (bright green dashed) shows a more predominant peak than that of the full catalog $\overline{\xi}_{160}$ (solid black). Both means show roughly the same peak position, and width. The vertical lines are 1σ uncertainties of the mean. Bottom Panel- Residual $\overline{\xi}_{75} - \overline{\xi}_{160}$ as function of scale.

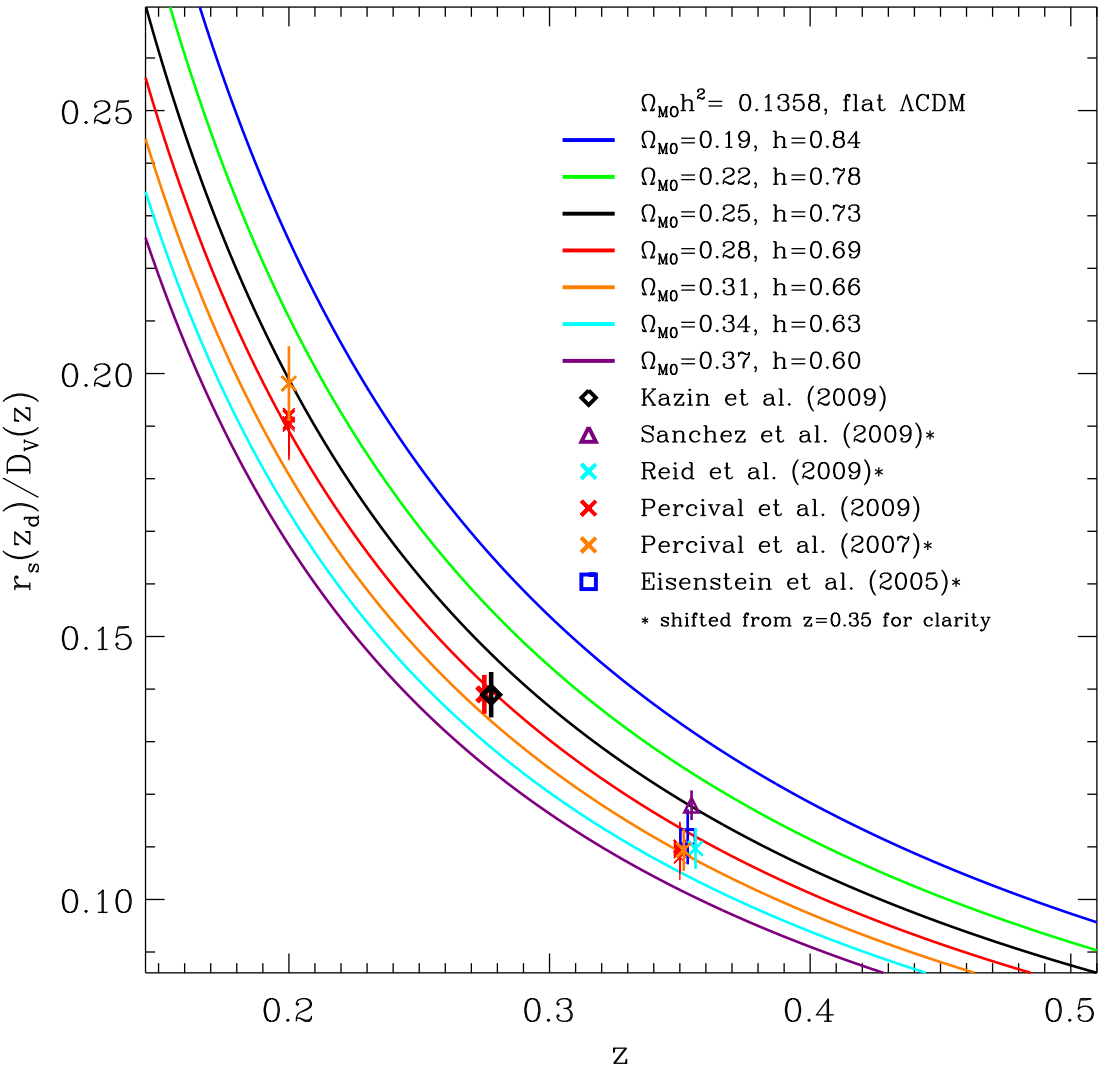

Fig. 11.— $r_s(z_d)/D_V(z)$ Result: We obtain $r_s/D_V(z = 0.278)$ of 0.1389 ± 0.0043 (black diamond; 1σ uncertainty) in good agreement with the $z = 0.275$ result presented by Percival et al. (2009) (red crosses). Other results (Sanchez et al. 2009: purple triangles, Reid et al. 2009: cyan crosses, Percival et al. 2007: orange crosses; Eisenstein et al. 2005: blue square) are indicated. These points are not all independent as they use similar samples. The solid lines show predictions of various flat ΛCDM cosmologies constraining $\Omega_{M0}h^2 = 0.1358$ and varying Ω_{M0} and h, where the top (blue) line is $\Omega_{M0} = 0.19$, $h = 0.84$ and the bottom (purple) line is $\Omega_{M0} = 0.37$, $h = 0.60$. Intermediate steps shown are $\Omega_{M0} = 0.22, 0.25, 0.28, 0.31, 0.34$. Our result clearly agrees with ranges $\Omega_{M0} = [0.25, 0.34]$ and $h = [0.63, 0.73]$.

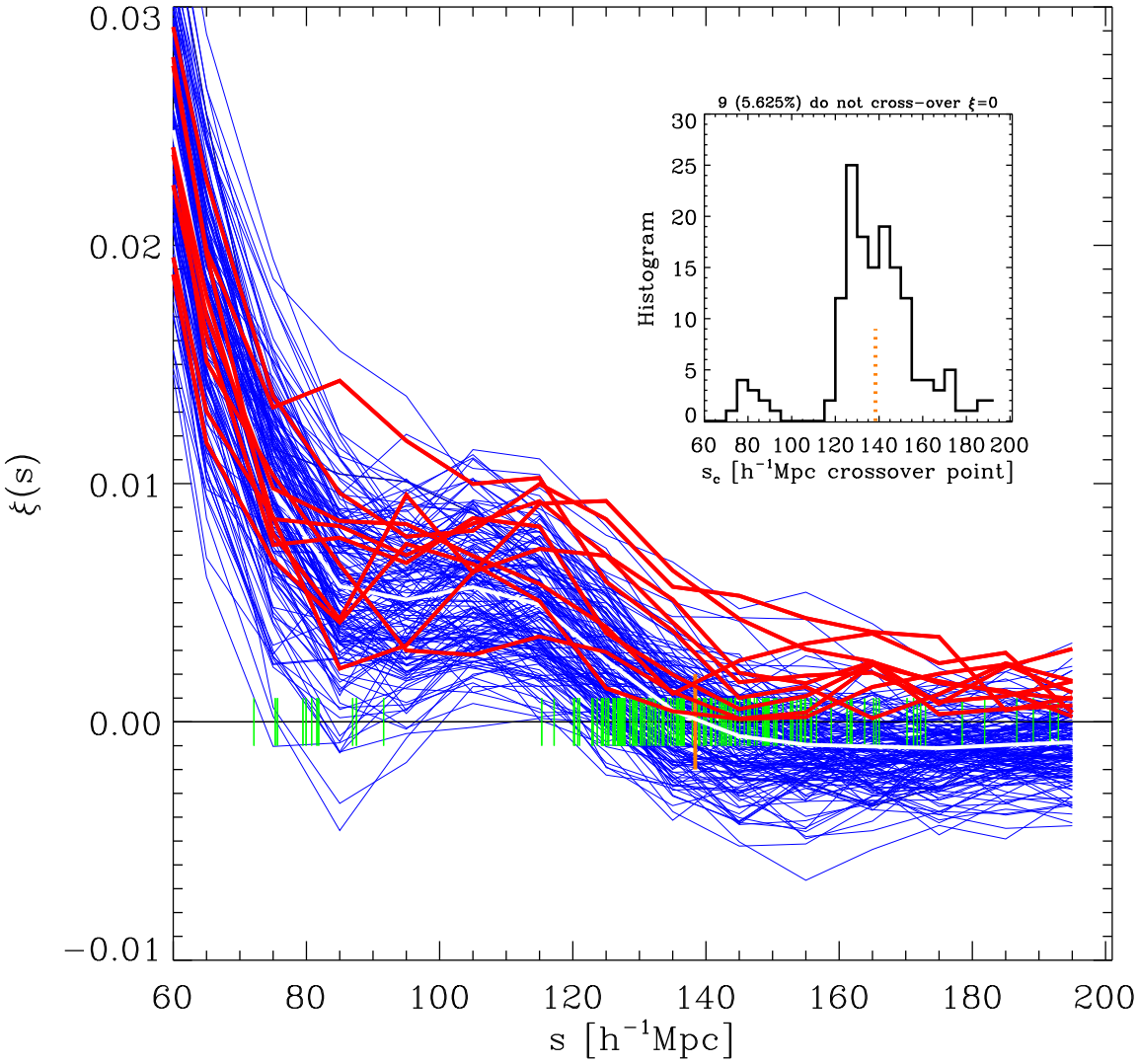

Fig. 12.— $\xi(s_c) = 0$: Here we show all 160 DR7-Dim mock realizations. Those with a crossover point s_c (crossover marked in vertical green lines) are blue solid lines, and the $\sim 6\%$ without are in thick red. The mean value is the solid white line, and its $s_c \sim 140\,h^{-1}$Mpc is indicated by the vertical orange line. In the inset we show a histogram of all s_c values, where the dotted orange line is the mean value.

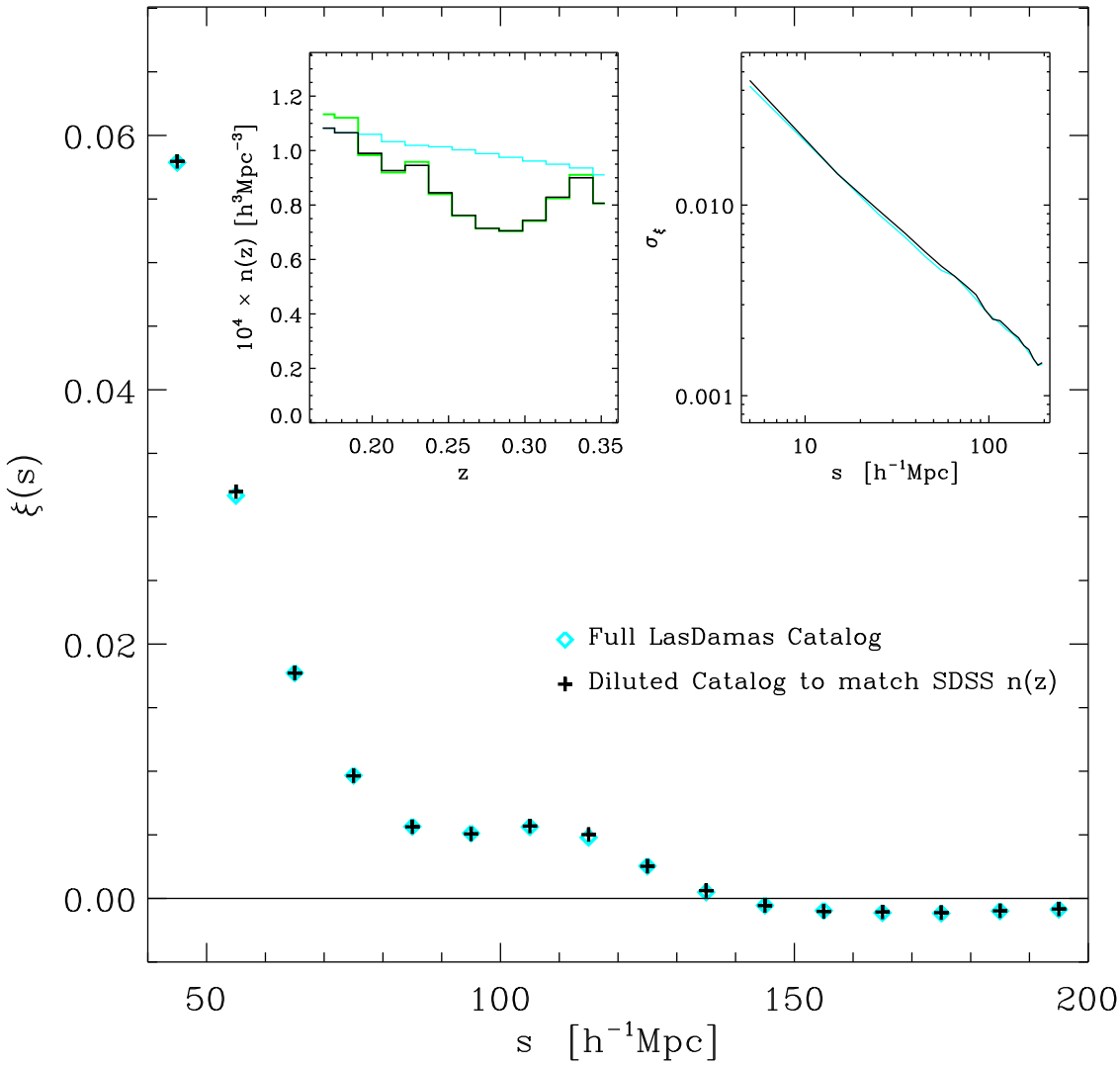

Fig. 13.— Radial selection effect on ξ and σ_ξ in DR7-Dim: Left inset shows $n(z)$ of LasDamas mock catalogs (light cyan line; $n(z)_1$) which has a slight negative slope, and that observed in the SDSS sample (thick green line). $n(z)_2$ (thin black line) is the reduced mock catalog that fits the SDSS $n(z)$. In the main plot the mean ξ over 160 mock DR7-Dim realizations for the different selection functions. Cyan diamonds are results of $n(z)_1$ and the black crosses are for $n(z)_2$. Both $\xi(s)$ are radially weighted with their respective $n(z)$, and are in very good agreement. Right inset shows the same for the r.m.s σ_ξ, indicating a slightly stronger variance for the latter sample, as it has slightly higher Poisson shot-noise.

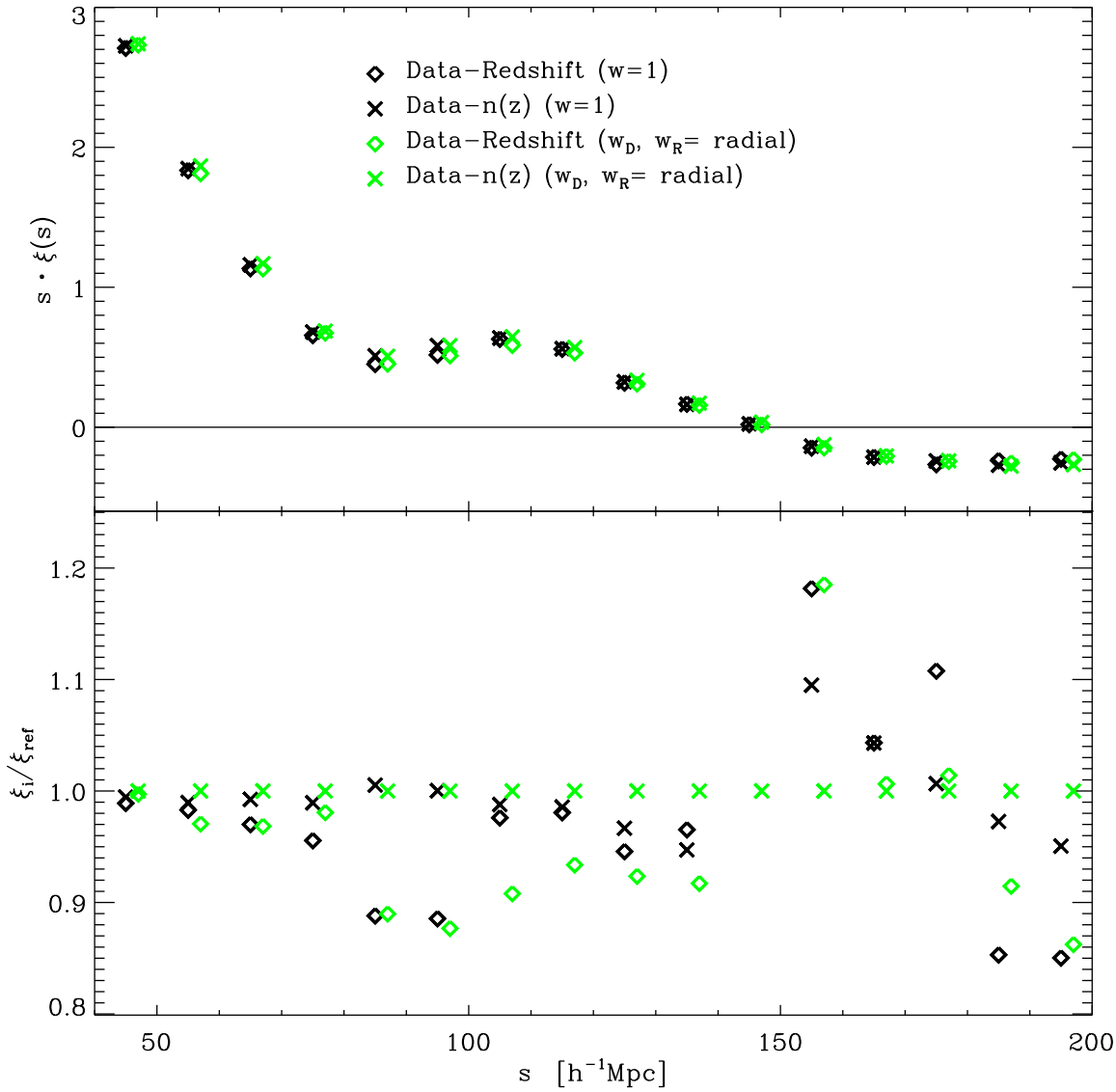

Fig. 14.— Random point distribution effects on ξ in DR7-Dim: Results for average over 8 mock realizations when using original mock $n(z)_1$ (Figure 13). In each realization we distribute radial distances to random points differently: diamonds- data point redshifts distributed randomly to random catalog (Data-Redshift), crosses- random $n(z)$ matches data but otherwise independent redshift distribution (Data-$n(z)$). Radial weighting schemes: Black- no weighting , Green- both data and random are weighted radially in same manner (shifted to right by $2\,h^{-1}$Mpc for clarity), Top Panel- To clarify differences we plot $s \cdot \xi(s)$. Bottom Panel- Ratio of each case ξ_i to weighted Data-$n(z)$ (ξ_{ref}). We clearly see that Data-Redshift (diamonds) yield the lowest result, hence diminishing the radial clustering mode. This is noticeable in range to $50 < s < 130\,h^{-1}$Mpc after which results agree with most of the other options. Weighting causes no effect in this case as the $n(z)$ used is close to volume-limited.

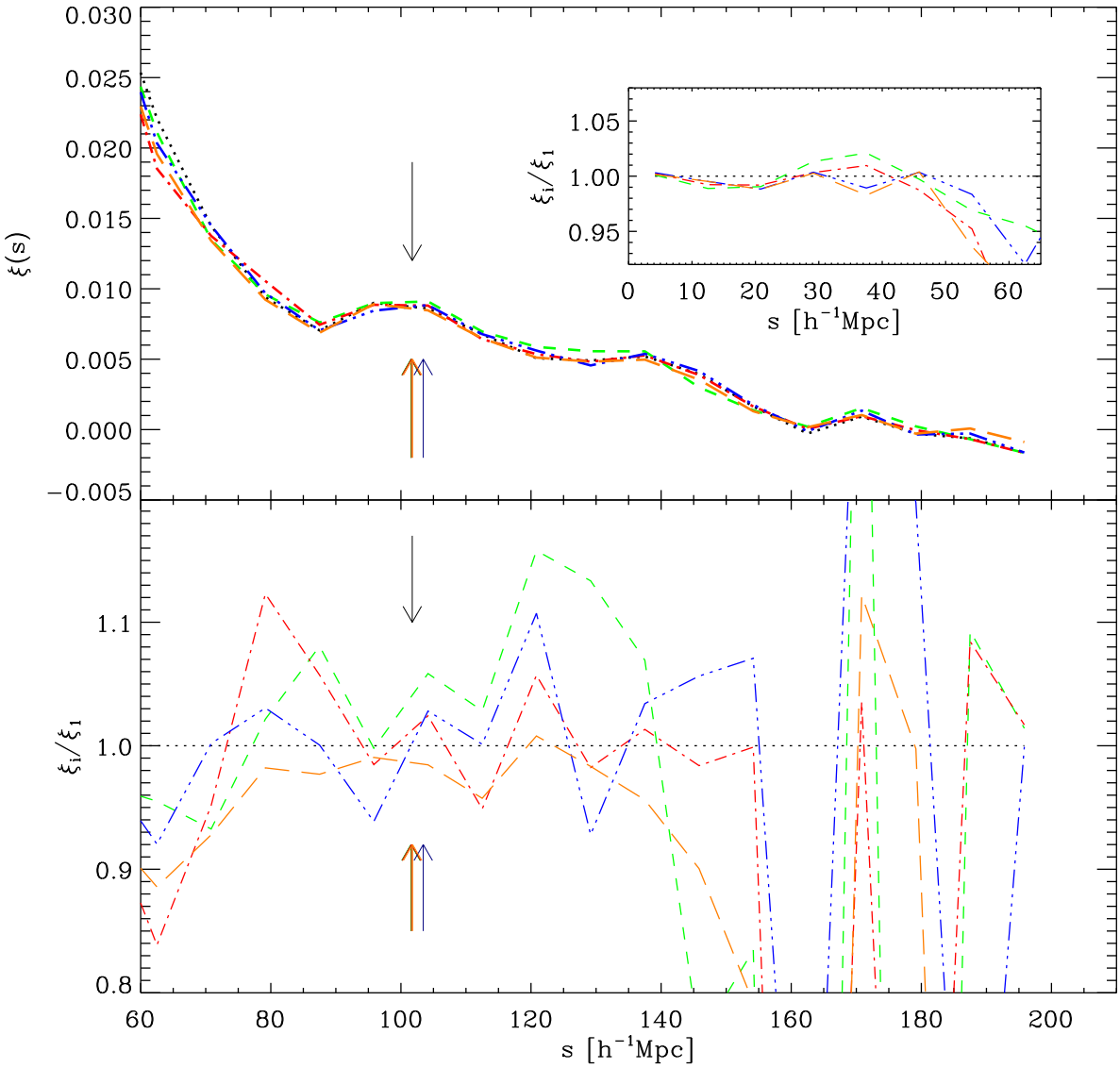

Fig. 15.— Random Shot-Noise: We run the same algorithm on the observed DR7-Dim, using five different random catalogs, each in turn. The ξ results are given in the top panel, where the five random sets are represented by: black-dotted, green-dashed, red-dot-short-dashed, orange-long-dashed, blue-triple-dotted-dashed lines. To facilitate differences the bottom panel shows the ratio of each $i^{\rm th}$ catalog ξ_i in respect to the first random catalog ξ_1. The inset has the same format as the bottom panel for smaller scales. The bottom arrows in both panels pin-point the peak position found, according to our algorithm. We see that four random catalogs result in same $s_{\rm p}$ within $1\,h^{-1}$Mpc where a fifth is $\sim 2\,h^{-1}$Mpc larger than the smallest obtained value. These results are for a ratio of random to data points of $r = {\rm N}_R/{\rm N}_D \sim 15.6$. To reduce this effect in our main analysis we use $r \sim 50$ (its result is marked by top arrow at $s_{\rm p} = 101.7\,h^{-1}$Mpc).

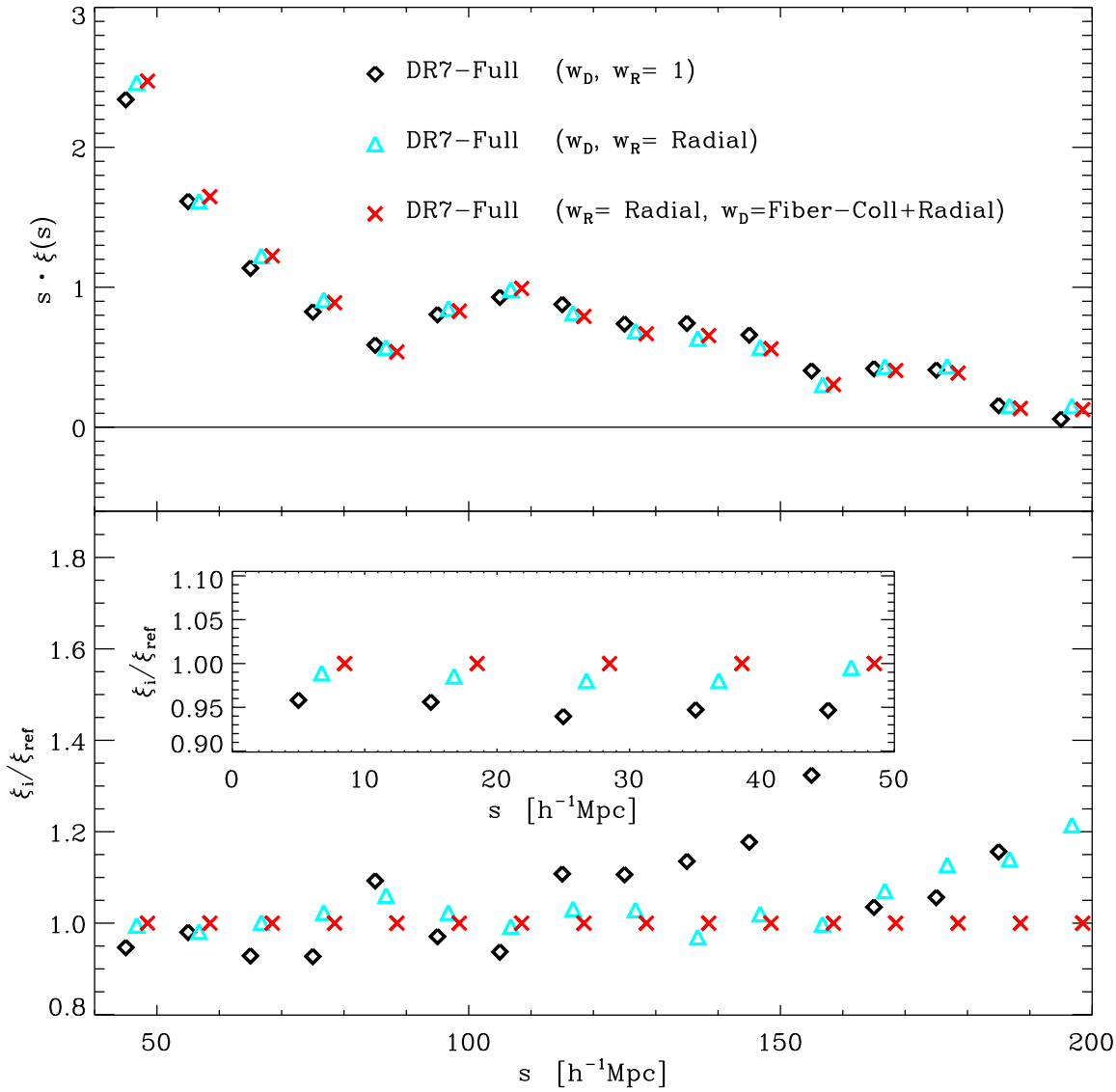

Fig. 16.— Weighting Effects on ξ: We compare different weighting schemes on the observed DR7-Full. w_D and w_R indicate the weight used on data and random points respectively, where black (diamonds) indicates no weight used ($w = 1$), cyan (triangles) indicates radial weight (Radial; shifted by $1.75\,h^{-1}$Mpc for clarity) and red (crosses) indicates radial+(fiber-collision correction) (Fiber-Collision+Radial; shifted by $35\,h^{-1}$Mpc). All options take into account sector completeness corrections. For clarity we plot in the top panel $s \cdot \xi(s)$ and on the bottom the ratio of each weighting option ξ_i over that chosen in this study (Fiber-Coll+Radial; ξ_{ref}). The inset is the same for smaller scales.

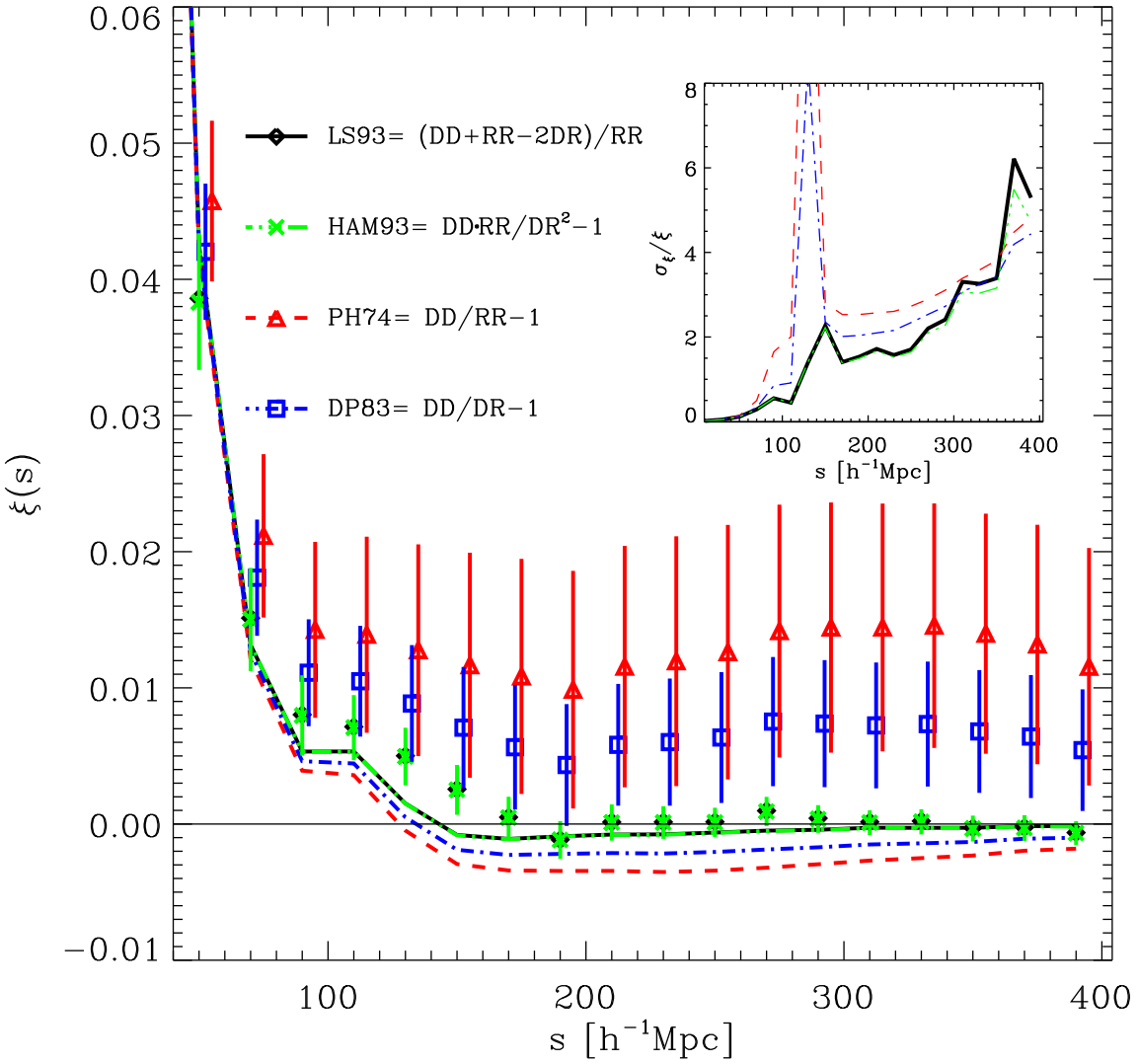

Fig. 17.— ξ Estimators at Large Scales of DR7-Dim: Comparing estimators proposed by LS93 (Landy & Szalay 1993; black), HAM93 (Hamilton 1993; green), DP83 (Davis & Peebles 1983; blue, symbols shifted for clarity) and PH74 (Peebles & Hauser 1974; red, symbols shifted). The lines are mock mean values, and the symbols are those observed (see legend). The uncertainty bars on the symbols are the r.m.s of the mocks. Although the observed results appear in reverse order than the mock mean, we verified by looking at particular mock realizations that some are in this order as well. Inset: shows the Noise to Signal σ_ξ/ξ. The strong spikes are due to sensitivity at the signal zero crossing. Refer to Figure 18 for $\sigma_\xi(s)$ of the different estimators. We find that LS93 and HAM93 agree very well on large scales, and asymptote quicker to zero than the other two. The ratio of random to data points used for SDSS is $r \sim 15.6$ and for mocks $r \sim 10$.

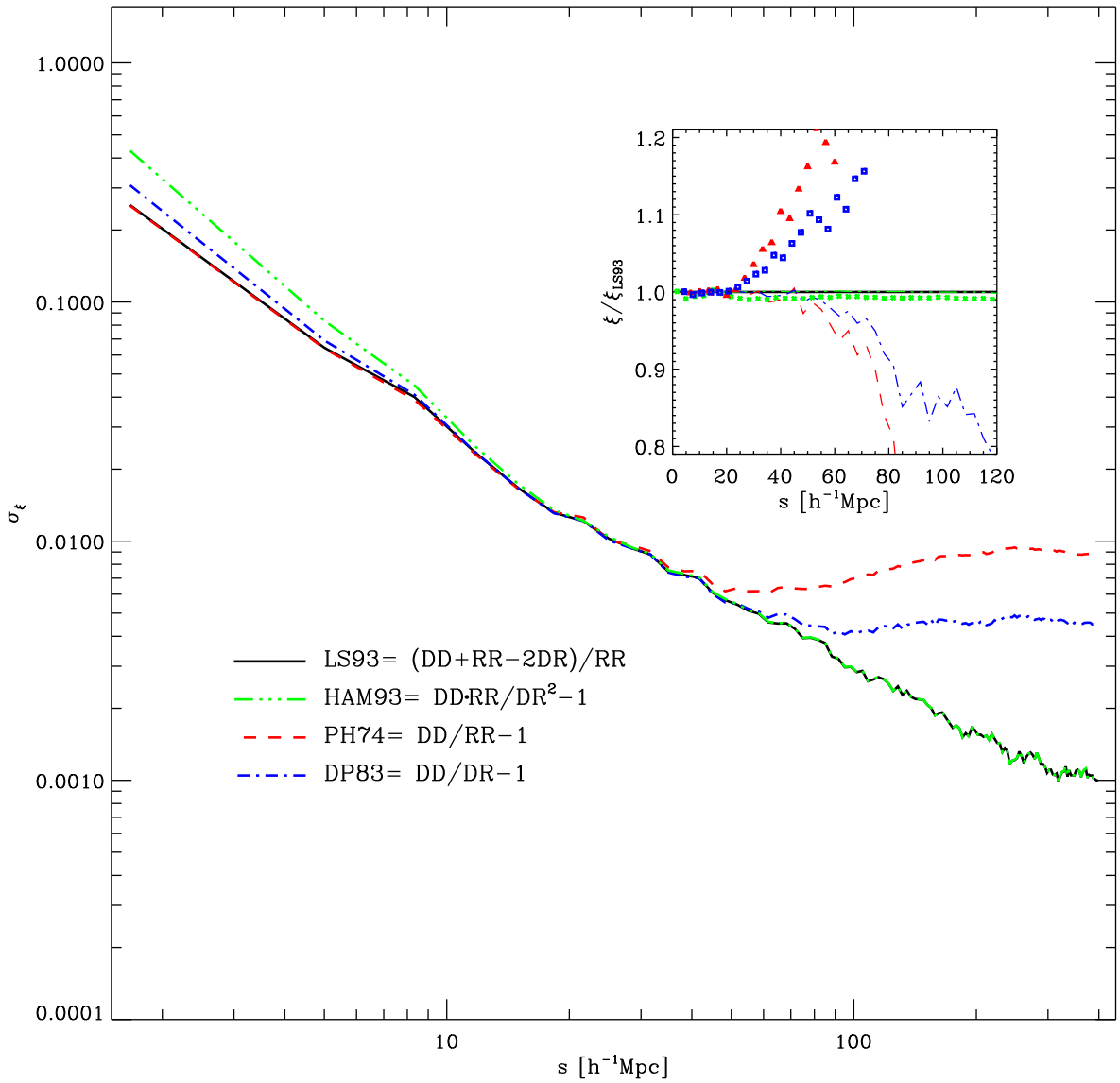

Fig. 18.— Estimator Variance: Main Plot: Same as Fig. 17 showing the r.m.s $\sigma_\xi = \sqrt{C_{ii}}$ where C_{ij} is the covariance of ξ. PH74 and DP83 show poor variance performance on large-scales. HAM93 agrees very well with LS93 on large scales but has a larger variation on $\lesssim 10\,h^{-1}$Mpc. Inset: Comparing estimators to Landy & Szalay (1993). The ratio of random to data points used for SDSS is $r \sim 15.6$ and for mocks $r \sim 10$.

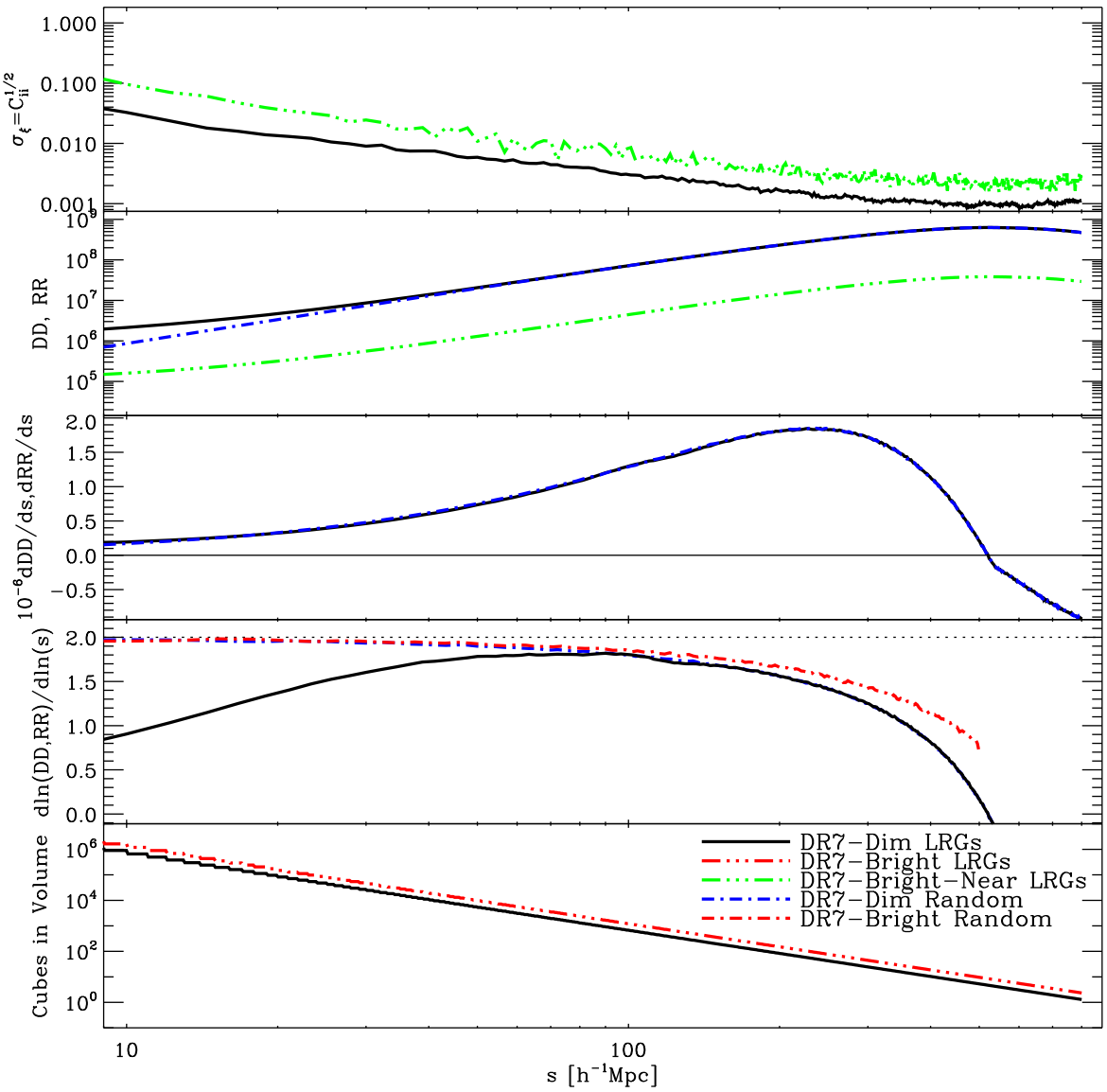

Fig. 19.— LRG and random point statistics up to $s = 800\,h^{-1}$Mpc. a) signal r.m.s σ_ξ for DR7-Dim (black solid line) and DR7-Bright-Near (green triple-dot-dashed). b) Normalized data-data pairs DD for DR7-Dim and DR7-Bright-Near (same notation) and random-random of DR7-Dim (blue dot-dashed) c) Same as (b) only derivative dx/ds. where x =DD,RR. d) Examining boundary effects in pair counting. A periodic box would yield $dln(RR)/dln(s) = 2$. Hence deviations from 2 (dotted line) are effects of boundary. e) Counting the number of cubes of length s fit in the survey volume, for DR7-Dim (solid black) DR7-Bright (red triple-dotted-dashed).

Table 1. SDSS LRG Samples

Sample	# of LRGs	$z_{\rm min}$	$z_{\rm max}$	$\langle z \rangle$	$M_{g,\rm min}$	$M_{g,\rm max}$	$\langle M_g \rangle$	Area ($\rm deg^{-2}$)	Volume (h^{-3} $\rm Gpc^3$)	Density (10^{-5} h^3 $\rm Mpc^{-3}$)
DR3[a]	47,063	0.16	0.47	0.327	−23.2	−21.2	−21.70	3,807	0.722	6.50
DR7-Full	105,831	0.16	0.47	0.324	−23.2	−21.2	−21.72	7,908	1.58	6.70
DR7-Dim	61,899	0.16	0.36	0.278	−23.2	−21.2	−21.65	7,189	0.66	9.40
DR7-Bright	30,272	0.16	0.44	0.338	−23.2	−21.8	−22.02	7,189	1.19	2.54
DR7-Bright-Near	16,473	0.16	0.36	0.284	−23.2	−21.8	−22.01	7,189	0.66	2.50
DR7-Bright-Far	13,799	0.36	0.44	0.402	−23.2	−21.8	−22.03	7,189	0.53	2.60
DR7-Bright2	32,861	0.16	0.36	0.282	−22.6	−21.6	−21.85	7,189	0.66	5.00

[a]Calculations of DR3 performed using $\Omega_{M0} = 0.3$.