DRAFT: October 3, 2002
Preprint typeset using LaTeX style emulateapj v. 5/12/01

Fermilab

FERMILAB-Pub-02/360-A January 2003

TESTING THE HALO MODEL AGAINST THE SDSS PHOTOMETRIC SURVEY

RYAN SCRANTON
Department of Astronomy and Astrophysics, University of Chicago, Chicago, IL 60637 USA
NASA/Fermilab Astrophysics Center, P.O. Box 500, Batavia, IL 60510 USA
scranton@oddjob.uchicago.edu
Draft: October 3, 2002

ABSTRACT

We present halo model predictions for the expected angular clustering and associated errors from the completed Sloan Digital Sky Survey (SDSS) photometric galaxy sample. These results are used to constrain halo model parameters under the assumption of a fixed ΛCDM cosmology using standard Fisher matrix techniques. Given the ability of the five-color SDSS photometry to separate galaxies into sub-populations by intrinsic color, we also use extensions of the standard halo model formalism to calculate the expected clustering of red and blue galaxy sub-populations as a further test of the galaxy evolution included in the semi-analytic methods for populating dark matter halos with galaxies. The extremely small sample variance and Poisson errors from the completed SDSS survey should result in very impressive constraints ($\sim 1-10\%$) on the halo model parameters for a simple magnitude-limited sample and should provide an extremely useful check on the behavior of current and future N-body simulations and semi-analytic techniques. We also show that similar constraints are possible using a narrow selection function, as would be possible using photometric redshifts, without making linear assumptions regarding the evolution of the underlying power spectra. In both cases, we explore the effects of uncertainty in the selection function on the resulting constraints and the degeneracies between various combinations of parameters.

Subject headings: large scale structure; cosmology; galaxies:halos; galaxies:evolution

1. INTRODUCTION

In the standard picture of structure formation, initial perturbations in the dark matter density collapse into halos (White & Rees, 1978; White & Frenk, 1991) in a hierarchical manner, starting at small mass scales and moving to larger masses over time. Simulations of this process have found that these halos have a self-similar shapes (Navarro, Frenk & White, 1996, NFW, hereafter; Moore et al., 1998), provided that one allows for some difference in central density for halos as a function of mass (Navarro, Frenk & White, 1997). Following this distribution of dark matter, we expect the baryonic matter to fall into these halos, cool and eventually form galaxies. This process can be simulated as well and the results can likewise be modeled by relatively simple semi-analytic methods (Kauffmann et al., 1999; Somerville & Primack, 1999; Benson et al., 2000). In addition to determining where galaxies form within a dark matter halo, these methods can also give estimates of the morphology, color, and star formation rates for galaxies, under certain assumptions.

The development of these prescriptions for the distribution of dark matter and galaxies has allowed for the calculation both the real-space two-point functions (Sheth & Jain, 1997; Jing et al., 1998; Peacock & Smith 2000) and power spectra (Seljak, 2000; Scoccimarro et al., 2000; Ma & Fry, 2000), generating predictions for galaxy clustering statistics in both the linear and nonlinear regimes which are both physically well-motivated and quite simple. These treatments can be further generalized to accommodate different galaxy sub-population clustering (Scranton, 2002; S02, hereafter), allowing for more detailed calculations.

On the data side, the next generation of galaxy surveys, in particular the Sloan Digital Sky Survey (SDSS; York et al. 2000; Gunn et al. 1998; Fukugita et al. 1996) have recently begun producing the anticipated large, rich galaxy catalogs. The initial galaxy clustering measurements (Zehavi et al., 2001; Connolly et al., 2002; Scranton et al., 2002; Gaztanaga 2001) have demonstrated not only the remarkable quality of data possible with a fully digital large area survey but also the enormous statistical power which will be available from such a combination of large area and depth of redshift. At the same time, the constraints on cosmological parameters (Ω_M, Ω_Λ, σ_8, etc.) have been improved by recent measurements of the CMB (cf. Pryke et al. (2001)) and large scale structure measurements (cf. Tegmark et al. (2001)) to the extent that we can reasonably consider the details of galaxy clustering and evolution in the context of a fixed cosmology.

The simultaneous development of these powerful tools for exploring galaxy clustering and evolution leads one to consider the possibilities for testing the predictions from the halo model against the future prospects of the data. Toward this end, we present calculations of the expected constraints on the halo model parameters from measurements of the angular clustering of galaxies for the completed SDSS survey. Since we wish to test not only the ability of the halo model to predict general galaxy clustering but also the modeling of galaxy evolution, we extend the general calculation to include clustering of the expected red and blue sub-populations. The exploration of sub-population clustering has been done to a certain extent in the SDSS redshift survey (Zehavi et al., 2001). However, many of the halo model parameters we will be investigating affect the respective power spectra on small scales ($k > \sim 1h\mathrm{Mpc}^{-1}$), a region where spectroscopic surveys are plagued by redshift-space distortions and observational complications (e.g. the collisions between spectroscopic fibers in the SDSS survey). Although these effects can be mitigated to a certain extent by projecting the clustering along the line of sight, one can achieve similar effects by considering the angular clustering in the photometric catalog. This approach also offers the benefit of a much deeper look ($z \sim 0.3$) at the clustering for all galaxies than the SDSS main galaxy spectroscopic sample ($z \sim 0.1$) is capable of delivering.

From the standpoint of the halo model, the importance of constraining the halo model parameters with observations is two-fold. The various components of the halo model (mass

function, concentration, etc.) have all been parameterized in various ways and fit to the results of simulations which try to replicate the underlying physics as close as is feasible. Our approach mirrors those efforts, replacing the simulation input with expected measurements on the sky. From the standpoint of the halo model as a theoretical construct, this sort of constraint will provide an invaluable check on whether the parameterizations coming out of simulations reasonably correspond to the data on the sky. If so, then we have a powerful tool for analytic calculations and future insight into the development of large scale structure and galaxy evolution. If not, we can isolate those parts of the formalism have failed (determining that the disagreement between simulation outputs and the observations is due to a failure in the assignment of galaxies to halos rather than the halo mass function, for instance). This feeds into the second aspect of the constraints: improving future simulations. The simulations carry greater information about the underlying physics than is present in the parameterizations that feed into the halo model calculations. Replacing simulation constraints on the halo model parameters with ones taken from the data should indicate which aspects of the physics that are included in the simulations are necessary to generate the observed clustering and evolution and what might be ignored.

In §2, we review the basics of the halo model and the augmentations necessary to calculate sub-population power spectra, establishing the fiducial set of model parameters we will use for all our calculations. With this laid out, §3 discusses the two types of selection functions we will use to project the three-dimensional galaxy clustering onto the sky. §4 briefly discusses the formalism for calculating the angular correlations as well as the corresponding covariance matrices and the additional information needed to account for uncertainties in the selection function. These angular correlations are converted into angular biases in §5, with their own attendant covariance matrix. These are then fed into the Fisher matrix formalism described in §6 to produce constraints on each of the parameters. Finally, §7 discusses the results of these calculations, the expected errors and degeneracies on each of the halo model parameters.

2. THE FIDUCIAL MODEL

For the purpose of calculating the power spectra that will feed into our angular correlations, there are two broad classes of parameters we will consider (we leave the details of how the parameters combine to produce power spectra to Appendix A). First, we have the general halo model parameters which describe the overall dark matter halo profile and biasing. These parameters have been measured from a number of N-body simulations and we will adopt conventional values.

The second set of parameters describes the halo occupation density: the abundance and distribution of galaxies (and galaxy sub-populations) in the halos. As mentioned previously, a variety of semi-analytic techniques have been applied to N-body simulations. For the purpose of our calculations, we will use the outputs of the GIF simulations (Kauffmann et al., 1999) which have been generated using the SDSS magnitudes. These particular implementations of the GIF methodology have not been the subject of extensive inquiry in the literature, giving us more flexibility to explore the parameterization possibilities. Further, given the relatively light computational load of calculating eventual constraints on these parameters, we will allow ourselves a generous parameter space, larger, in fact, than could easily be constrained by an actual measurement of the angular clustering for red and blue galaxies given simple computational methods. However, by performing this seemingly over-zealous calculation, we can determine which of the parameters are well-constrained by the angular correlations and which might be fixed without significantly affecting the errors on the other more sensitive parameters.

2.1. *Dark Matter Halo Parameters*

Beginning with the general halo parameters, the fundamental unit of the halo model is the halo profile. This can be parameterized along the lines of the profile derived by NFW,

$$\rho(r) = \frac{\rho_S}{(r/r_S)^{-\alpha}(1+r/r_S)^{(3+\alpha)}} \tag{1}$$

where r_S is the universal scale radius and $\rho_S = 2^{3+\alpha}\rho(r_S)$. We can replace r_S by a concentration, $c \equiv r_v/r_S$, where r_v is the virial radius. This radius is defined as the radius encompassing a region within which the fractional overdensity of the halo Δ_V (Eke, et al., 1996) scales as

$$\Delta_V(z) = 18\pi^2(\Omega_M(z)^{-0.55}), \tag{2}$$

where $\Omega_M(z)$ is the matter density relative to the critical density for a given redshift,

$$\Omega_M(z) = \left[1 + \frac{1-\Omega_M}{(1+z)^3\Omega_M}\right]^{-1}, \tag{3}$$

Ω_M is the matter density today relative to the critical density, and we have chosen a ΛCDM cosmology where $\Omega_M + \Omega_\Lambda = 1$. The concentration is a weak function of halo mass ($c \equiv c_0(M/M_*)^\beta$, where $c_0 \sim O(10)$ and $\beta \sim -O(10^{-1})$). The traditional NFW profile gives $\alpha = -1$, while the Moore profile has $\alpha = -3/2$. We will use $\alpha = -1.3$ for the calculations in this paper, but the general results are largely insensitive to the choice of α. Bullock et al. (2001) gives $c_0 = 9$ for a pure NFW profile; using Peacock & Smith's relation, $c_0 \approx 4.5$ for a Moore profile. Since we are using an intermediate value of α, we choose $c_0 = 6$ and $\beta = -0.15$ for all the calculations in this paper. In order to generate power spectra at a variety of redshifts, we also scale the concentration as $c_0 \sim (1+z)^{-1}$ for a given redshift z, as in Bullock et al. (2001).

Once we know how the mass in a halo is distributed, we need to know how many halos of a given mass we expect to find, i.e. the halo mass function. Traditionally, this mass function is expressed in terms of a function $f(\nu)$,

$$\frac{dn}{dM}dM = \frac{\bar{\rho}}{M}f(\nu)d\nu, \tag{4}$$

where ν relates the minimum spherical over-density that has collapsed at a given redshift (δ_c) and the rms spherical fluctuations containing mass M ($\sigma(M)$) as

$$\nu \equiv \left(\frac{\delta_c}{\sigma(M)}\right)^2, \tag{5}$$

This can be generalized for an arbitrary redshift by taking the forms from Navarro, Frenk and White (1997),

$$\delta_c(z) = \frac{3}{20}(12\pi)^{2/3}(\Omega_M(z)^{0.0055}) \tag{6}$$

and scaling $\sigma(M)$ as $\sigma(M,z) \equiv \sigma(M,0)D(z)$ where $D(z)$ is the linear growth factor for a given redshift z normalized to unity at $z = 0$. We define M_* as the mass corresponding to $\nu = 1$. The functional form for $f(\nu)$ is traditionally given by the Press-Schechter function (1974). This form tends to over-predict the

number of halos below M_*, so we use the form found from simulations by Sheth and Tormen (1997),

$$\nu f(\nu) \sim (1+\nu'^{-p})\nu'^{1/2}e^{-\nu'/2}, \quad (7)$$

where $\nu' = a\nu$, $a = 0.707$ and $p = 0.3$. This gives us a total of five general halo model parameters we might constrain with angular measurements: α, c_0, β, a and p.

2.2. *Galaxy Parameters*

As described in Appendix A, to calculate the galaxy power spectrum, we need to know the mean number of galaxies for a given halo mass ($\langle N\rangle(M)$) and the second moment of the galaxy distribution ($\langle N(N-1)\rangle(M)$). Following the example in S02, we will generate different power spectra for red and blue galaxies by using different $\langle N\rangle$ relations for each of the sub-populations as well as changing the distribution of the galaxies within each halo. Before settling on the specific values of the respective parameters, we need to establish some necessary formalism.

Figure 1 shows the measurements of $\langle N\rangle(M)$ taken from the GIF simulations for red and blue galaxies. As in S02, we can parameterize the $\langle N\rangle(M)$ relations as modified power laws. In the case of red galaxies, the power law is cut-off by a lower mass limit ($M_{\rm R0}$),

$$\langle N\rangle_{\rm R}(M) = \left(\frac{M}{M_{\rm R}}\right)^{\gamma_{\rm R}} e^{-(M_{\rm R0}/M)^{1/2}}. \quad (8)$$

For the blue galaxies, the modification is a bit more involved, including a Gaussian term to account for low-mass halos with a single blue galaxy. Thus, our $\langle N\rangle_{\rm B}(M)$ relation is

$$\langle N\rangle_{\rm B}(M) = \left(\frac{M}{M_{\rm B}}\right)^{\gamma_{\rm B}} + Ae^{-A_0(\log(M)-M_{\rm Bs})^2}. \quad (9)$$

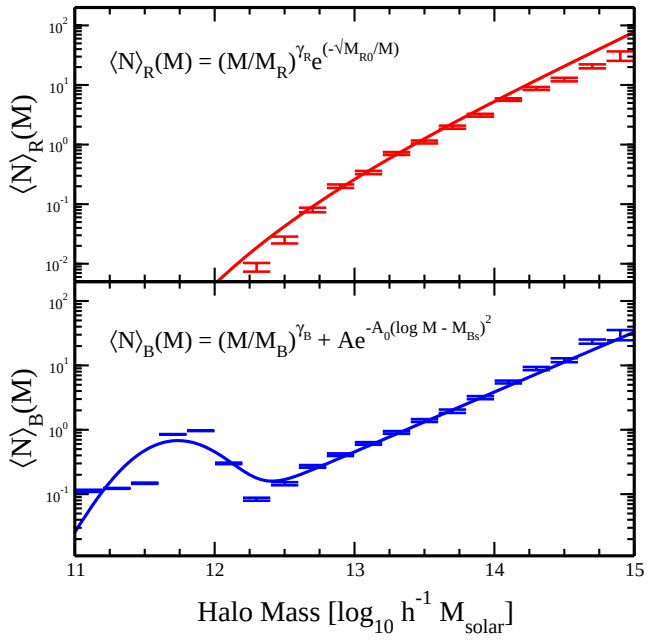

FIG. 1.— $\langle N\rangle(M)$ for red and blue galaxies as measured at $z = 0.27$ in the GIF simulations. The solid lines indicate the values of $\langle N\rangle_{\rm R}$ and $\langle N\rangle_{\rm B}$ using the forms in Equations 8 and 9 and parameter values in Table 1. The errorbars in both cases represent the Poisson errors in each mass bin.

TABLE 1
FIDUCIAL MODEL PARAMETERS

Class	Parameter	Value
Dark Matter Halo	α	-1.3
	c_0	6
	β	0.15
	a	0.707
	p	0.3
Galaxy HOD	μ	50.6
	η	3
	$M_{\rm R}$	$1.8\times10^{13}h^{-1}M_\odot$
	$\gamma_{\rm R}$	1.1
	$M_{\rm R0}$	$4.9\times10^{12}h^{-1}M_\odot$
	$M_{\rm B}$	$2.34\times10^{13}h^{-1}M_\odot$
	$\gamma_{\rm B}$	0.93
	A	0.65
	A_0	6.6
	$M_{\rm Bs}$	11.73
	M_α	12.1
	σ_α	0.27

NOTES.—Galaxy Halo Occupation Density (HOD) parameters taken from GIF simulations at $z = 0.27$. Red galaxies taken to have $g'-i' > 0.85$ in rest frame colors.

As mentioned above, we must also adopt a form for the second moment of the distributions, $\langle N(N-1)\rangle(M)$. In S02, we followed the results found by Scoccimarro et al. (2001) which related $\langle N(N-1)\rangle$ to $\langle N\rangle$ as $\langle N(N-1)\rangle(M) = \alpha_M^2(M)\langle N\rangle^2(M)$, where

$$\alpha_M(M) = \begin{cases} 1 & M > 10^{13}h^{-1}M_\odot \\ \log\left(\sqrt{M/10^{11}h^{-1}M_\odot}\right) & M < 10^{13}h^{-1}M_\odot \end{cases} \quad (10)$$

This gives the galaxies the Poisson distribution at large mass and sub-Poisson distribution observed in simulations. Since the mass dependence of α_M can significantly affect the small scale power, we would like to include this effect in our parameter constraints. However, to make the form easier to incorporate into our later Fisher matrix formalism, we use a sigmoid function,

$$\alpha_M(M) = \left[1+e^{(\log(M)-M_\alpha)/\sigma_\alpha}\right]^{-1}, \quad (11)$$

where M_α and σ_α are chosen to match the behavior in the GIF simulations (fitting to match the behavior in Equation 10 results in values of 12 and 0.45 for M_α and σ_α, respectively).

While Equations 8 and 9 are sufficient to describe the expected number of red and blue galaxies in a halo of given mass, they do not suffice to match the observation that red and blue galaxies have different radial distributions within a given halo, with red galaxies tending to populate the halo center and blue galaxies the outer regions. To generate different distributions for our red and blue galaxies within each halo, we follow the method outlined in S02, using profiles of the form given in Equation 1 as the distribution functions for the red and blue galaxies within each halo. In principle, one could construct distribution functions for each of the galaxy sub-populations from observations of their relative abundances at a number of radii. However, by assuming a form for the distributions, the power-law indices and relative normalizations of these distri-

butions are effectively determined by considering two parameters: the ratio of blue to red galaxies at large halo radii (η) and the inverse ratio at small radii (μ). Since these ratios relate to the number of galaxies, rather than the mass assigned to the galaxies, we have to transform these quantities through our $\langle N \rangle$ relations from Equations 8 and 9 in order to maintain mass conservation within the halo. Thus, we define η' and μ' as

$$\eta' = \left[\eta - Ae^{-A_0(\log(M_R)-M_{Bs})^2}\right]^{1/\gamma_B} \frac{M_B}{M_R} \quad (12)$$
$$\mu' = \left[\mu\left(1 + Ae^{-A_0(\log(M_B)-M_{Bs})^2}\right)\right]^{1/\gamma_R} \frac{M_R}{M_B}.$$

For large radii, the profiles will both scale as $\rho \sim r^{-3}$, so η' sets the relative normalization directly:

$$\rho_{SR} = \frac{1}{\eta'+1}\rho_S \quad (13)$$
$$\rho_{SB} = \frac{\eta'}{\eta'+1}\rho_S,$$

The choice of radius (r_i) for the measurement of μ is somewhat arbitrary, so we follow S02 in setting $r_i c/r_v = 0.1$, (where c is the halo concentration and r_v is the virial radius) for our calculations. This gives us the difference in the power law indices for the red and blue galaxies ($\Delta\alpha$),

$$\Delta\alpha \equiv \alpha_B - \alpha_R = \frac{\log(\mu'\eta')}{\log(1+r_i c/r_v) - \log(r_i c/r_v)}. \quad (14)$$

With this relation between α_R and α_B in hand, we can perform a simple search over values of α_R to find the sub-profiles that combine to closest match an overall profile with a given value of α. Since we know that $\Delta\alpha$ must be positive, this relation guarantees a flatter distribution of blue galaxies in the center of halos relative to red galaxies and Equation 13 produces relatively more blue galaxies in the outer regions.

With this formalism in place, the only elements missing are actual values for the parameters in Equations 8, 9, 11, and 12. Since we need the $\langle N \rangle$ relations to calibrate μ, η, M_α and σ_α, we begin with $\langle N \rangle_R$ and $\langle N \rangle_B$.

Typically, the $\langle N \rangle$ relations are determined from simulations using very wide or open ended magnitude cuts. However, as shown in Figure 2, we can see that the shape of the $\langle N \rangle$ relations can change quite dramatically if we consider only a narrow range in apparent magnitude (like those described later in §3), particularly the location of the Gaussian component. As a consequence of this complication, any constraints made on the parameters in $\langle N \rangle_R(M)$ and $\langle N \rangle_B(M)$ from a given magnitude cut will not accurately describe the complete number-mass relationships. This situation can be salvaged to some extent however, based upon the fact that shape of the number-mass surface does not appear to vary strongly with redshift (the entire surface does shift to fainter magnitudes with increasing redshift, as one would expect for apparent magnitudes). Thus, with a volume limited sample, one could in principle constrain the entire surface with a series of magnitude limited measurements.

Leaving aside these concerns, we can fit the parameters in Equations 8 and 9 for the magnitude cut in §4 ($20 \leq r' \leq 21$) for a typical peak redshift in the selection functions at $z \sim 0.3$. The results of these fits are shown in Table 1; a comparison of the fits to the outputs of the simulations and a discussion of the color cut used are given in Appendix B. The parameters listed show a significant difference from previous similar fits (e.g Scoccimarro et al. (2001)). This is not surprising, however, since our fits are made at higher redshift, resulting in a

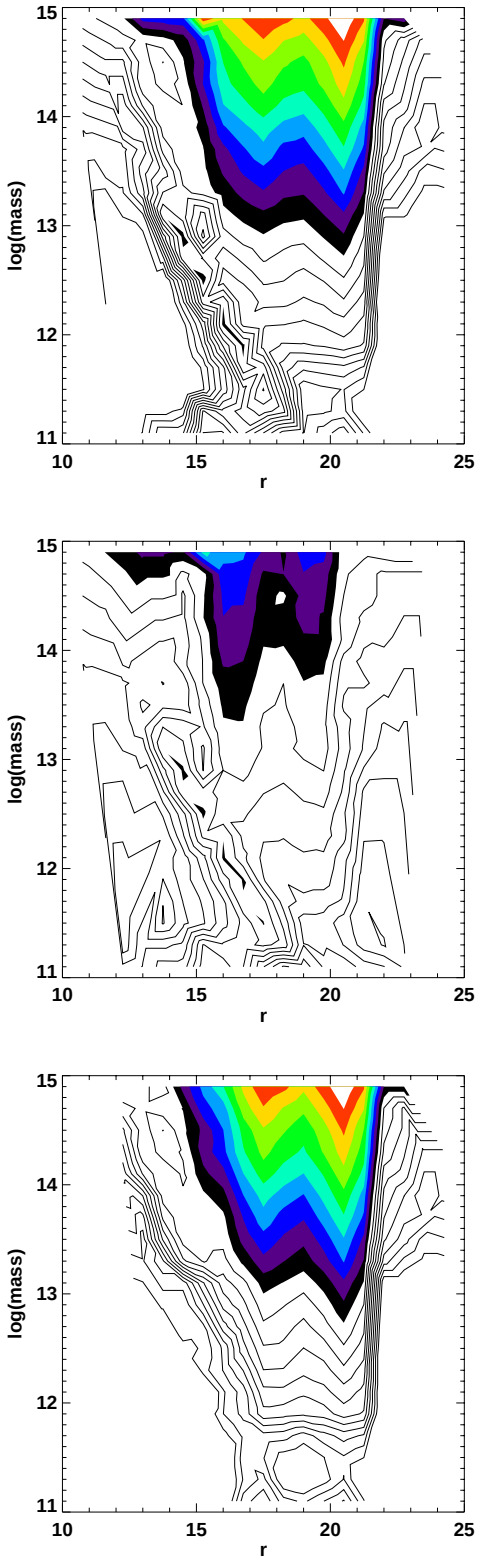

FIG. 2.— $\langle N \rangle$ as a function of halo virial mass and r' magnitude from the GIF simulations at $z = 0.06$. From top to bottom, the panels give the surfaces for all galaxies, blue galaxies (rest frame $g' - i' < 0.85$) and red galaxies (rest frame $g' - i' > 0.85$), respectively. Contours correspond to a change in 0.25 for $\log(\langle N \rangle)$. Filled contours indicate $\log(\langle N \rangle) > 0$ and wire-frame contours indicate $\log(\langle N \rangle) < 0$.

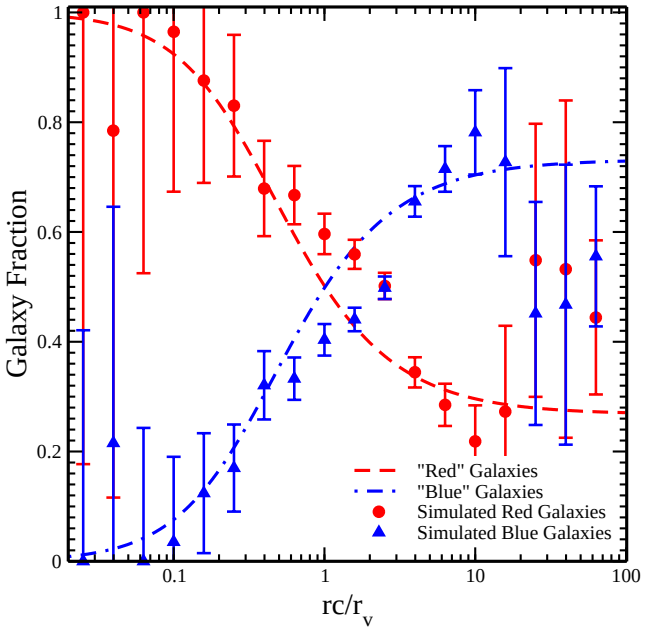

FIG. 3.— Comparison of fraction of red and blue galaxies as a function of rc/r_v. The lines give the distributions predicted from the sub-population distributions discussed in §2.2 and the points show the distribution of galaxies from the GIF simulations. For the simulated data, red galaxies were taken to have $g' - i' > 0.85$ in the rest frame (see Appendix B for more details). Error bars represent the Poisson error for all of the galaxies in a given radial bin, scaled appropriately for the fractional comparison.

correspondingly higher mass scale for all of the $\langle N \rangle$ relations. Likewise, the narrow magnitude cut leads to a relative suppression of the $\langle N \rangle$ relations at smaller halo mass giving larger values of γ_B than have been reported previously. It is interesting, however, that the combination of magnitude and redshift did not significantly change the $\langle N(N-1) \rangle$ mass scale, allowing us to use similar values in Equation 11 as those matching the form of Equation 10. At this redshift, we also observe a higher value of μ in the inner regions of the halo, although the value of η is largely unchanged from the $z = 0$ value in S02. To test this set of parameters, we can compare the distribution of red and blue galaxies in the simulation to that which we would predict from our model. As we can see in Figure 3, this combination of parameters reconstruct the halo galaxy distribution observed in the simulations reasonably well.

Putting the parameters from Table 1 into our prescription for calculating the power spectra leads to the curves shown in Figure 4 for $z = 0.3$. As with the $z = 0$ power spectra in S02, there is significant biasing in the red galaxy sample and anti-biasing in the blue galaxies. Likewise, there is a similar (but weaker) large k break in the blue power spectrum. The power spectrum for the whole galaxy sample shows a slight break around $k \sim 10h\mathrm{Mpc}^{-1}$ which is not seen in the $z = 0$ case or in low redshift observations (cf. Hamilton et al., 2000). At this small scale, the power spectra are dominated by the smallest mass halos which, in turn, are almost exclusively populated by blue galaxies, particularly at higher redshifts. This suggests that we should not be surprised to find a break in the power law behavior of $w_{XX}(\theta)$ for a photometric redshift-selected galaxy sample.

3. REDSHIFT DISTRIBUTION

TABLE 2

REDSHIFT DISTRIBUTION PARAMETERS FOR A SIMPLE MAGNITUDE CUT AND NARROW WINDOW DISTRIBUTION.

Distribution	Galaxy Type	a	z_0	b
Magnitude Cut	Early	2.6	0.37	3.3
	Intermediate	2.0	0.37	2.45
	Late	0.98	0.4	2.28
	All	1.28	0.42	2.94
Narrow Window	Early	12	0.35	15
	Intermediate	12	0.3	15
	Late	12	0.3	15
	All	12	0.32	15

To project the three dimensional power spectrum onto the sky we need to know the redshift distributions of the red and blue galaxies in our sample. In order to separate the galaxies into red and blue classes, we would need to have photometric redshift information available on each galaxy, meaning that we could, in principle, choose the galaxy sample to match any selection function we desired within the bounds of the survey limits. Alternatively, we could use the redshift distribution resulting from a simple magnitude cut. The magnitude cut results in a somewhat broad distribution of redshifts, so we will use the arbitrary redshift distribution to explore the effects of a much narrower redshift window. In either case, however, we are limited by the ability of the photometric redshift calculations to cleanly separate our galaxy sample into red and blue types, which, in turn, is limited by the photometric errors at a given magnitude. Based on early results from the photometric redshift work on the SDSS, the practical limit for reliable redshifts is $r' \approx 21$.

For the case of a simple magnitude cut, we can follow the method used by Dodelson et al (2001) in modifying the redshift

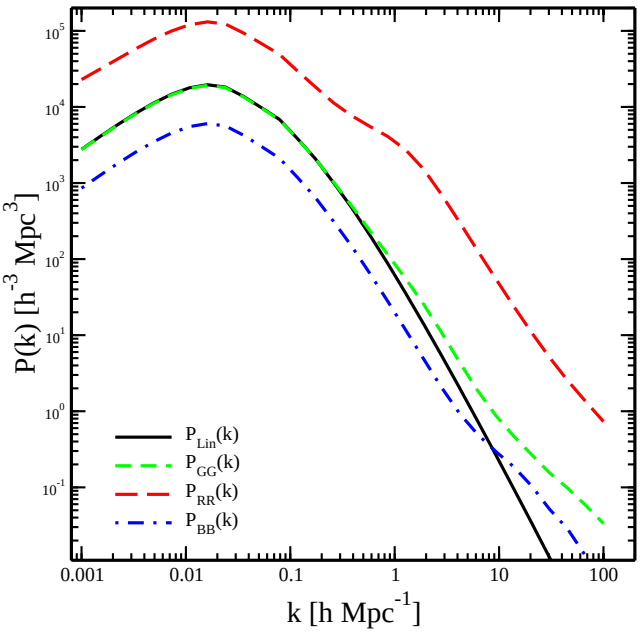

FIG. 4.— Power spectra at $z = 0.3$ for red ($P_{RR}(k)$) and blue ($P_{BB}(k)$) galaxies compared to the linear dark matter ($P_{Lin}(k)$) and galaxy ($P_{GG}(k)$) power spectra.

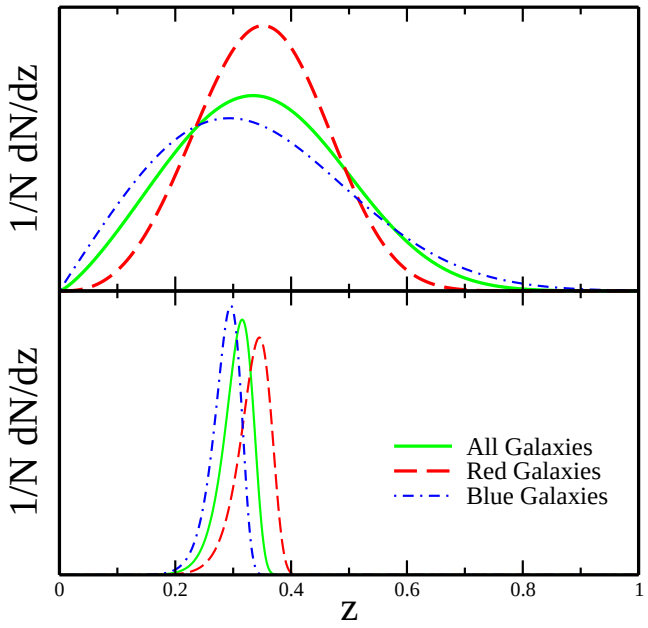

FIG. 5.— The top panel gives the normalized redshift distributions for all galaxies with $20 \leq r' \leq 21$ and the red and blue sub-populations, as modified for SDSS colors from the CNOC2 survey. The bottom panel gives an example of narrow redshift distributions possible for a sample of galaxies using photometric redshifts.

distributions found from the CNOC2 survey (Lin et al, 2000) to match the SDSS filters. The redshift distributions for red and blue galaxies (taken from Lin's early type and the combination of intermediate and late types, respectively) can be seen in the left panel of Figure 5. Since these distribution functions are taken from the morphological types rather than the color cut mentioned in Table 1 and Appendix B, we do not expect these distributions to exactly match those found in the final SDSS data, but they should give us a reasonable approximation. The shape of these distributions can be well approximated by a function of the form

$$\frac{dN_X}{dz} \sim z^a \exp\left(-(z/z_0)^b\right). \quad (15)$$

The fits to the parameters in this function for red and blue galaxies with $20 \leq r' \leq 21$ are given in Table 2. We also can use this form to specify the shape of the narrow redshift distribution, adopting the second set of parameters in Table 2, resulting in the redshift distributions shown in the right panel of Figure 5. By design, these artificial distributions have roughly the same peak in redshift as the magnitude cut distributions, which should make the eventual comparison independent of evolution effects in the power spectrum. In addition, they all avoid the region around $z \sim 0.4$ as much as possible. This is particularly important due to the limitations of the photometric redshifts available in with the SDSS filters; there is a degeneracy in the colors for early type galaxies near a redshift of 0.4 resulting in large uncertainty in the photometric redshift for these objects.

4. CALCULATING $w(\theta)$ AND COVARIANCE MATRICES

4.1. *Limber's Equation*

With the power spectra and selection functions in hand, we can calculate the expected angular correlations for the red and blue galaxies ($w_{RR}(\theta)$ and $w_{BB}(\theta)$, respectively) using Limber's equation:

$$w_{XX}(\theta) = \frac{1}{2\pi}\int dk k \int d\chi F_X(\chi) P_{XX}(k,\chi) J_0(k\theta\chi) \quad (16)$$

where χ is the comoving angular diameter distance and J_0 is the Bessel function. We could also consider the angular cross-correlation between red and blue galaxies, but since we will be re-casting the angular correlations in terms of relative biases in §5, the cross-correlation will not yield any additional information. The normalized redshift distributions, $F_R(\chi)$ and $F_B(\chi)$, are given by

$$F_X(\chi) = \left[\frac{1}{N_X}\frac{dN_X}{dz}\frac{dz}{d\chi}\right]^2, \quad (17)$$

where

$$\frac{dz}{d\chi} = H_0\left[\Omega_M(1+z)^3 + \Omega_K(1+z)^2 + \Omega_\Lambda\right]^{1/2} \quad (18)$$

for a given Hubble constant (H_0), matter density (Ω_M), cosmological constant (Ω_Λ) and curvature ($\Omega_K = 1-\Omega_M-\Omega_\Lambda$). As mentioned previously, we will use a flat ΛCDM model ($\Omega_M = 0.3$, $\Omega_\Lambda = 0.7, h = 0.7$) and normalize the linear power spectrum such that $\sigma_8 = 0.9$ and $n_S = 1$.

We can simplify Equation 16 by assuming a linear scaling relation for the power spectrum, $P_{XX}(k,\chi) = D^2(\chi)P_{XX}(k)$, where we choose $P_{XX}(k)$ to have the shape of the power spectrum at the peak redshift of $F_X(\chi)$. Making this substitution gives us

$$w_{XX}(\theta) = 4\pi^2 \int dk k P_{XX}(k) \int d\chi F_X(\chi) D^2(\chi) J_0(k\theta\chi). \quad (19)$$

The results of performing this calculation for the three power spectra and the two sets of selection function are shown in Figure 6. In addition to the assumption of linear power spectrum scaling, we also truncate our integral over wavenumber at $k = 100h\mathrm{Mpc}^{-1}$; this does not change the calculated values of $w_{XX}(\theta)$ due to the flat shape of the kernel at small $k\theta$ and allows us to avoid any complication of the power spectrum shape due to baryon concentration in the innermost regions of the halo.

The price we pay for the assumption of linear power spectrum scaling is the neglect of any change in the shape of the

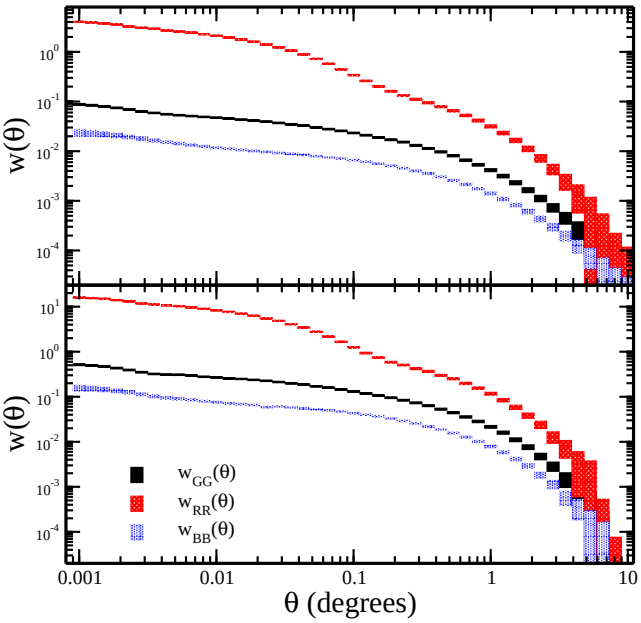

FIG. 6.— Angular correlations for all galaxies ($w_{GG}(\theta)$), red galaxies ($w_{RR}(\theta)$) and blue galaxies ($w_{BB}(\theta)$). The upper panel shows the angular correlations for the magnitude-based selection function and the lower for the photometric redshift selection function.

power spectrum at large k as a function of redshift due to non-linear effects. For the photometric redshift selection functions, this is not a serious problem, but the width of the magnitude-based selection function gives one pause, particularly when considering the very small expected errors (§4.2). However, given the nature of the surfaces in Figure 2, it is clear that any changes in the shape of the power spectrum we might see will be a reflection of the changing $\langle N\rangle$ relations as a function of redshift. This makes our choice of fiducial model for $\langle N\rangle_{\rm R}$ and $\langle N\rangle_{\rm B}$ somewhat nebulous. We can see the effects of calculating $w_{\rm XX}(\theta)$ with and without the linear approximation for the two selection functions in Figure 7. As expected, the photometric redshift selection function shows little to no effect even when we calculate the power spectrum using the full $\langle N\rangle$ and $\langle N(N-1)\rangle$ surfaces. For the magnitude selection selection function, the results are identical when we hold the $\langle N\rangle$ relations the same, but show strong deviations if we use the $\langle N\rangle$ and $\langle N(N-1)\rangle$ surfaces to calculate $w_{\rm XX}(\theta)$ without the linear approximation. Fortunately, as we will see in §7, the equivalent constraints obtained with the photometric redshift selection function make this concern moot.

4.2. *Covariance Matrices*

Since we will need to account for the fact that the angular bins in our measurement are going to be highly correlated, we must deal with the full covariance matrix and not just the expected errors on each bin. For the expected Gaussian and non-Gaussian sample variance covariance matrices, we follow the prescription given in Eisenstein & Zaldarriaga (2001) (as well as similar treatments in Cooray & Hu (2001) and Scoccimarro, et al. (1999)). The sample variance ($\mathbf{C}_S(\theta,\theta')$) consists of two parts, a component dependent only on the two-point angular correlation function (the Gaussian covariance) and a second piece which is a function of the four-point angular correlation (the non-Gaussian covariance):

$$\begin{aligned}\mathbf{C}_S(\theta,\theta') &\equiv \left\langle (w(\theta)-\hat{w}(\theta))(w(\theta')-\hat{w}(\theta'))\right\rangle \\ &= \mathbf{C}_G(\theta,\theta')+\mathbf{C}_{NG}(\theta,\theta').\end{aligned} \quad (20)$$

In order to make the calculation simpler, we can re-write Equation 19 in terms of the angular power spectrum ($P_{\rm XX}(K)$),

$$w_{\rm XX}(\theta) = \int \frac{KdK}{2\pi} P_{\rm XX}(K) J_0(K\theta), \quad (21)$$

where $P_{\rm XX}(K)$ is given in angular wavenumber space ($K \equiv k\chi$) and

$$P_{\rm XX}(K) = \frac{1}{K}\int dk P_{\rm XX}(k) F_{\rm X}(\chi) = \int \frac{d\chi F_{\rm X}(\chi)}{\chi^2} P_{\rm XX}(k). \quad (22)$$

Using this formulation, we can write the Gaussian covariance matrix ($\mathbf{C}_G(\theta,\theta')$) as

$$\mathbf{C}_G(\theta,\theta') = \frac{1}{\pi A_\Omega}\int dK K P^2_{\rm XX}(K) J_0(K\theta) J_0(K\theta'), \quad (23)$$

where the area of the survey (A_Ω) is π steradians in the case of the SDSS.

To calculate the non-Gaussian component, we need to generate an estimate of the trispectrum for our halo model. Fortunately, as Cooray and Hu (2001) indicate, the majority of the non-Gaussian covariance can be accounted for by merely calculating the single halo contribution of the trispectrum. This term, which is independent of configuration under the assumption of spherical halos, is given by

$$\begin{aligned}&T_{4\rm X}(k_1,k_2,k_3,k_4) = \quad (24)\\ &\frac{\bar{\rho}}{\bar{n}^4}\int_0^\infty f(\nu)\frac{\langle N\rangle^{(4)}}{M(\nu)}|y(k_1,M)||y(k_2,M)||y(k_3,M)||y(k_4,M)|d\nu,\end{aligned}$$

where $\langle N\rangle^{(4)} \equiv \alpha_M^4 \langle N\rangle^4$ under the assumptions in Equation 11. As with the case of the Gaussian component, we need to project the trispectrum into angular wavenumber space,

$$T_{4\rm X}(K_1,K_2,K_3,K_4) = \int d\chi \frac{F^2_{\rm XX}(\chi)}{\chi^6} T_{4\rm X}(k_1,k_2,k_3,k_4). \quad (25)$$

Finally, the angular trispectrum needs to be averaged over an annulus in angular wavenumber space,

$$\bar{T}_{\rm XX}(K_1,K_2) = \int \frac{d^2K_1}{A_r}\int \frac{d^2K_2}{A_r} T_{4\rm X}(K_1,K_1,K_2,K_2), \quad (26)$$

where A_r is the area of the annulus. In the limit of narrow binning, we can approximate $\bar{T}_{\rm XX}(K_1,K_2)$ by calculating

$$T_{\rm XX}(k_1,k_2) = \frac{\bar{\rho}}{\bar{n}^4}\int_0^\infty f(\nu)\frac{\langle N\rangle^{(4)}}{M(\nu)} y^2(k_1,M) y^2(k_2,M) d\nu, \quad (27)$$

and appropriately transforming into angular wavenumber space. With this in hand, we can calculate the non-Gaussian component of the sample variance using

$$\begin{aligned}&\mathbf{C}_{NG}(\theta,\theta') = \quad (28)\\ &\frac{1}{4\pi^2 A_\Omega}\int dK K \int dK' K' \bar{T}_{\rm XX}(K,K') J_0(K\theta) J_0(K'\theta').\end{aligned}$$

As the final component of the statistical error, we can add a Poisson term ($\mathbf{C}_P(\theta,\theta')$) to the diagonal elements,

$$\mathbf{C}_P(\theta,\theta') = \frac{A_\Omega}{N^2\delta\Omega}\delta_{\theta,\theta'}, \quad (29)$$

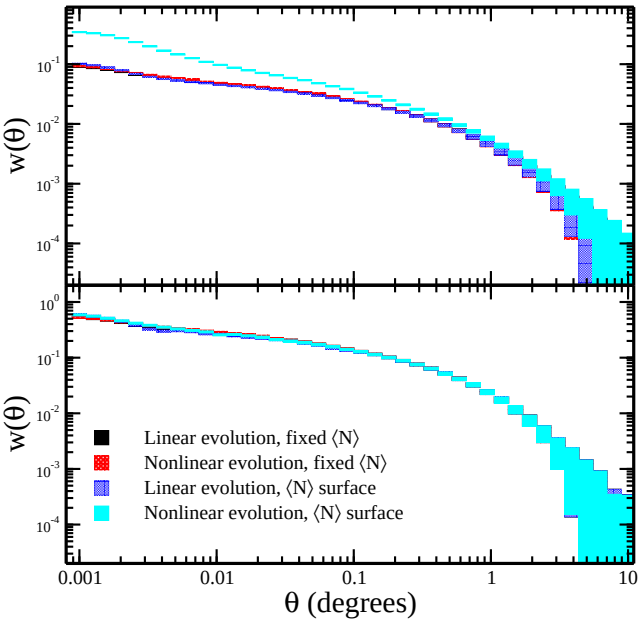

FIG. 7.— Angular correlations for all galaxies ($w_{\rm GG}(\theta)$) under a number of assumptions for power spectrum evolution. The black and red curves show the angular correlations (and associated errors) which have been calculated using the $\langle N\rangle_{\rm R}$, $\langle N\rangle_{\rm B}$ and α_M relations given by Equations 8, 9 and 11 (respectively) and the parameter values in Table 1. The black curve assumes the linear power spectrum evolution in Equation 19, while the red curve calculates $w_{\rm GG}(\theta)$ using Equation 16. The blue and cyan curves take their $\langle N\rangle$ and $\langle N(N-1)\rangle$ relations taken directly from surfaces like those shown in Figure 2 (shifted in magnitude appropriately for a given redshift) to calculate the power spectrum under linear and non-linear assumptions of Equations 19 and 16, respectively. The upper panel shows the angular correlations for the magnitude-based selection function and the lower for the photometric redshift selection function.

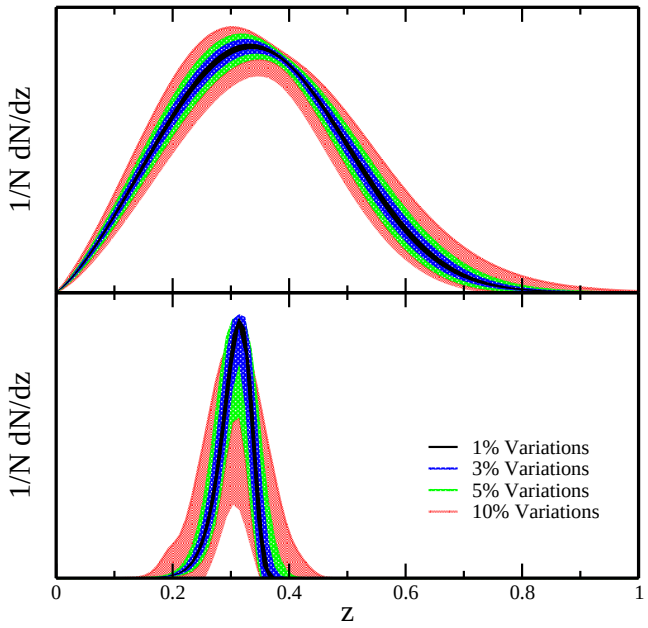

FIG. 8.— Variations in the redshift distribution function errors. The top panel gives the variations for the magnitude-based selection function and the bottom for the photometric redshift selection function.

where N is the total number of galaxies, $\delta\Omega$ is the area of the angular bin and $\delta_{\theta,\theta'}$ is the standard Kronecker delta. For the full photometric survey, the SDSS will contain on order 200 million objects. We can scale this appropriately for the $20 \leq r' \leq 21$ magnitude cut from §3, using 50 million galaxies for the full sample and half that for each of the sub-samples. For the photometric redshift selection function, we can combine it with the magnitude selection function, resulting in 8 million galaxies within the redshift and magnitude ranges.

In addition to the statistical errors due to sample variance and shot noise, we need to consider the errors in our calculation of the expected angular correlations due to uncertainty in the underlying redshift distribution. Indeed, in §7, we will see that, given the relatively large area observed and number of galaxies which will be in the final SDSS data set, the dominant source of error in our final constraints will come from the error in the redshift distributions. To model the errors in the selection function, we follow the treatment given in Dodelson et al. (2001). Since we do not have the exact errors for the parameter fits to the eventual redshift distributions, we will consider errors on those parameters of 1, 3, 5 and 10 percent. Drawing from 50,000 Monte Carlo realizations of the redshift distributions with these variations, we can calculate the covariance of the expected angular correlation functions ($\mathbf{C}_Z(\theta,\theta')$) and use this to approximate the covariance from the redshift distribution uncertainty. Figure 8 shows the average deviation from the magnitude cut redshift distribution for each of the parameter variation levels. Bringing these three pieces together gives us our final covariance matrix ($\mathbf{C}(\theta,\theta')$)

$$\mathbf{C} = \mathbf{C}_G + \mathbf{C}_{NG} + \mathbf{C}_P + \mathbf{C}_Z. \tag{30}$$

Figure 9 shows the error in $w_{\rm GG}(\theta)$ due to each term in the covariance matrix as a function of angle. In the absence of selection function errors, the errors on the very smallest angular scales are dominated by shot noise, giving way to Gaussian sample variance at larger angles. Non-Gaussian sample variance is small enough to ignore even at small angles. When the errors due to uncertainties in the selection function are included, they quickly dominate the statistical errors on small angular scales in all but the most modest error regime.

5. RELATIVE BIASES

In principle, the angular correlations we have calculated so far are sufficient to calculate the parameter constraints. However, the range of values that $w_{\rm XX}(\theta)$ takes on as a function of θ can be a problem for numerical derivatives. To avoid that, we can instead use the angular biases given by the ratios of the angular correlations,

$$\begin{aligned} b^2_{\rm RB}(\theta) &= \frac{w_{\rm RR}(\theta)}{w_{\rm BB}(\theta)} \\ b^2_{\rm RG}(\theta) &= \frac{w_{\rm RR}(\theta)}{w_{\rm GG}(\theta)} \\ b^2_{\rm BG}(\theta) &= \frac{w_{\rm BB}(\theta)}{w_{\rm GG}(\theta)}, \end{aligned} \tag{31}$$

where $w_{\rm GG}(\theta)$ is the angular correlation for all of the galaxies. As Figure 10 shows, this switch not only decreases the absolute range of we must consider, but also gives us more features in the curves to help determine the parameters. The price that we pay for this improvement is an additional step in the calculation of the covariance matrix we will use later in §6. Likewise, the larger amplitude $w_{\rm XX}(\theta)$'s for the photometric redshift based selection function (and commensurate larger values in the covariance matrix) results in proportionally larger errors on the resulting relative biases.

In order to properly determine the errors on the biases, we need to take into consideration the correlations between the angular bins in $w_{\rm XX}(\theta)$ as indicated by each correlation function's covariance matrix. To do this, we decompose the covariance

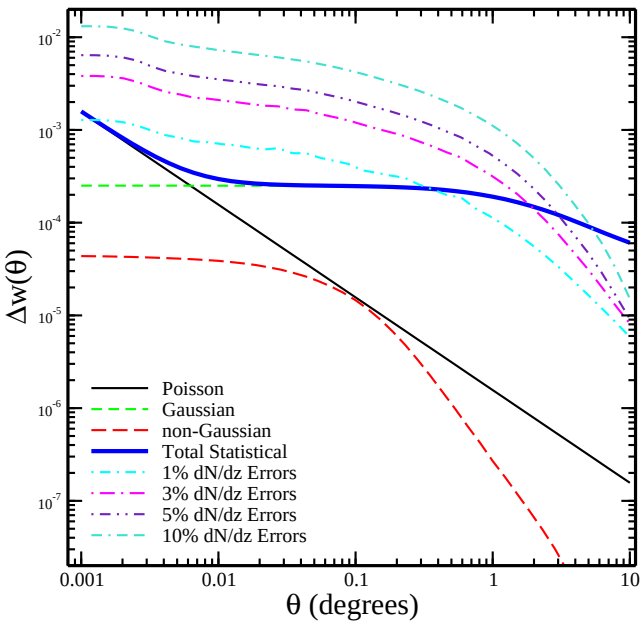

FIG. 9.— Statistical and selection function contributions to the error on $w_{\rm GG}(\theta)$ for the magnitude-based selection function. For the photometric redshift selection function, the results are similar for the Poisson, Gaussian and selection function errors. The contribution from non-Gaussian sample error is relatively stronger for the photometric selection function, but still remains small enough to safely ignore in the total covariance matrix.

matrix (**C**) into its eigenmodes, creating a basis (**R**) where each of the modes is independent and the variance on that mode is given by the associated eigenvalue (c_i). We can project $w_{\rm XX}(\theta)$ into that basis as

$$w_{\rm XX}(\theta)' = \mathbf{R}^T w_{\rm XX}(\theta), \tag{32}$$

where $\mathbf{R}^T$ is the transpose of **R**. Since we require that the covariance matrix be positive definite, we set any c_i which is negative due to numerical errors to zero and remove that mode from $\mathbf{r}'$. Within this basis, we can create a set of Monte Carlo realizations of each $w_{\rm XX}(\theta)'$ which, when transformed back into the angular basis, will have the correct covariance. Thus, the mean ratios of these realizations will give us the values for Equation 32 with the proper correlation between angular bins and the covariance between these realizations gives us the correct errors on our relative biases.

6. FISHER MATRIX CALCULATION

To estimate the expected errors on the parameters in Table 1, we can use the standard Fisher matrix formalism. Choosing a fiducial set of parameters leads to a reference angular bias, $\hat{b}^2(\theta)$. We can approximate the likelihood for some variation of the parameters as

$$L = \frac{1}{\left(2\pi C_B^D\right)^{N/2}} \exp\left[-\frac{1}{2}\delta(\theta)\mathbf{C}_B^{-1}(\theta,\theta')\delta(\theta')\right], \tag{33}$$

where N is the number of angular bins, $\mathbf{C}_B(\theta,\theta')$ is the covariance matrix for the angular biases from the Monte Carlo calculations in §5, C_B^D is the determinant of $\mathbf{C}_B(\theta,\theta')$ and $\delta(\theta) \equiv (b^2(\theta) - \hat{b}^2(\theta))$. The Fisher matrix is related to the likelihood function as

$$\mathbf{F}_{\alpha\beta} = -\left\langle \frac{\partial \ln L}{\partial x_\alpha \partial x_\beta} \right\rangle, \tag{34}$$

and $(\mathbf{F}^{-1})_{\alpha\beta}$ gives us the covariance between parameters α and β marginalized over all other parameters, while $(\mathbf{F}_{\alpha\beta})^{-1}$ gives us the covariance without marginalizing over all other parameters. Plugging our likelihood function from Equation 33 into Equation 34 gives us the Fisher matrix in terms of first derivatives,

$$\mathbf{F}_{\alpha\beta} = \frac{\partial b^2(\theta)}{\partial x_\alpha}\mathbf{C}_B^{-1}(\theta,\theta')\frac{\partial b^2(\theta')}{\partial x_\beta}. \tag{35}$$

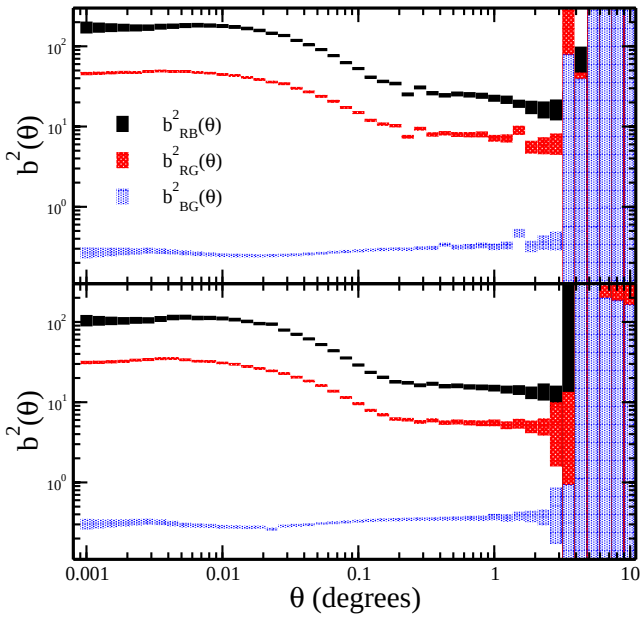

FIG. 10.— Relative angular biases $b^2_{\rm RB}(\theta)$, $b^2_{\rm RG}(\theta)$ and $b^2_{\rm BG}(\theta)$ for the magnitude-based selection function (upper panel) and the photometric redshift selection function (lower panel).

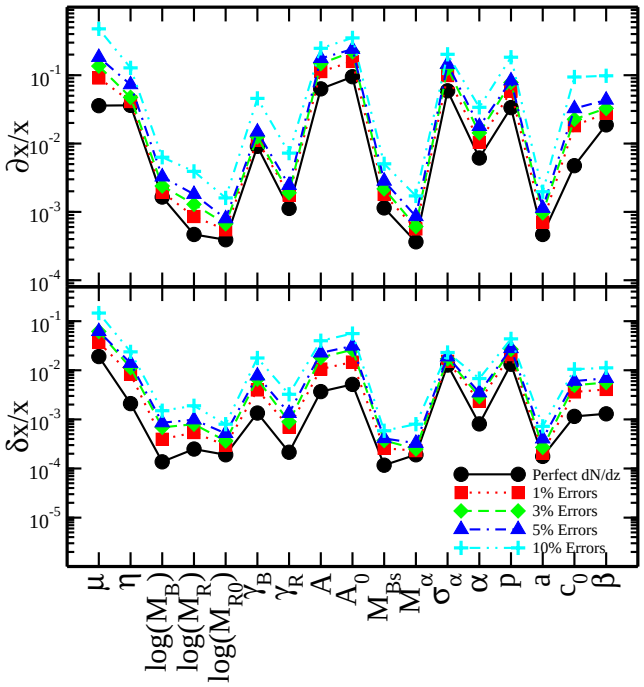

FIG. 11.— Marginalized and unmarginalized errors (∂x and δx, respectively) for each parameter in each of the selection function error regimes for the magnitude-based selection function.

In principle, only two of the relative biases given in Equation 32 would be necessary to constrain the halo model parameters. However, given the amplitude of the errors when the effects of errors in the selection function are included (particularly in the 10% case), we find a more stable solution when using all three relative biases. This makes our $b^2(\theta)$ a concatenation of $b^2_{\rm RB}(\theta)$, $b^2_{\rm RG}(\theta)$ and $b^2_{\rm BG}(\theta)$ and requires us to calculate a joint covariance matrix for all biases. In order to eliminate the presence of singular modes in this joint matrix, we only consider those angular bins from $b^2_{\rm RB}(\theta)$, $b^2_{\rm RG}(\theta)$ and $b^2_{\rm BG}(\theta)$ which have errors less than the amplitude of the bias in all three measurements, reducing our number of angular bins by about one third. The addition of the third relative bias does result in some degenerate modes in our covariance matrix, but a standard singular value decomposition routine can handle these modes adequately, resulting in numerically identical Fisher matrices for either 2 or 3 relative biases in the small selection function error cases.

In addition to these concerns, there is also the need to use a sufficient number of Monte Carlo realizations to ensure that the statistical noise from the realization (which goes as $\delta \sim 1/\sqrt{N}$ for N realizations) is sufficiently low to allow for an accurate numerical calculation of the derivatives in Equation 35. To meet this requirement, we used one million Monte Carlo realizations for each derivative calculation, achieving a statistical noise in the derivatives on order 0.1%. Given this level of precision, we calculated the derivatives using centered derivatives with

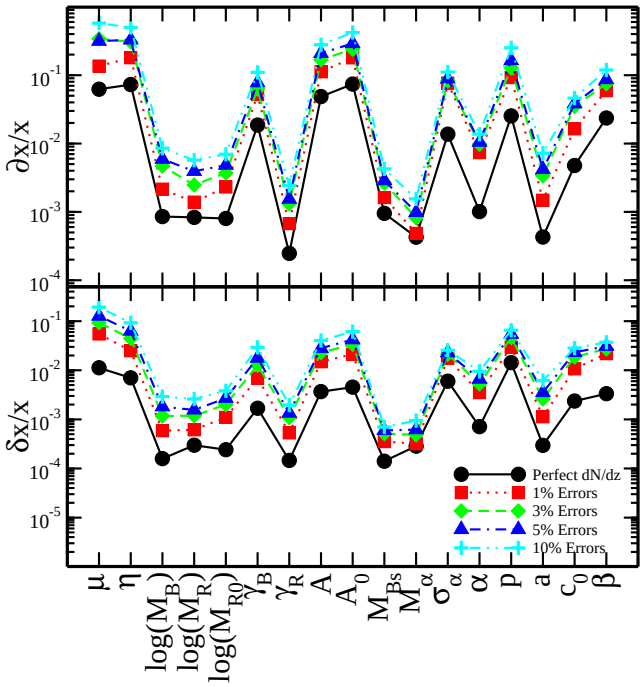

FIG. 12.— Same as Figure 11, but for the photometric redshift selection function.

a typical step size ($\Delta x/x \equiv \delta^{1/3}$) of 10% for each parameter. Given the much larger absolute value of $M_{\rm R}$, $M_{\rm B}$, and $M_{\rm R0}$ than the remainder of our parameters, we calculated the estimated constraint on the logarithm of each of these parameters, rather than the full value. Because of this transformation we reduced the step size for these parameters (as well as $M_{\rm Bs}$ and M_α) to 5%. The accuracy of the resulting Fisher matrix was verified by comparison with a Fisher matrix calculated from the estimates of the noise in each derivative, typically resulting in a noise on the diagonal of the Fisher matrix of less than half a percent.

7. RESULTS

As with any such calculation, there are essentially two questions to be addressed: what is the expected magnitude of the errors on each parameter and what are the expected degeneracies between the various parameters. We will take these questions in turn.

7.1. *Error Magnitudes*

Figures 11 and 12 give the fractional errors ($\partial x/x$ and $\delta x/x$) for each of the selection functions and selection function error regimes, where we take ∂x to be the marginalized error on parameter x, and δx to be the unmarginalized error:

$$(\partial x)^2 \equiv (\mathbf{F}^{-1})_{xx} \qquad (36)$$
$$(\delta x)^2 \equiv (\mathbf{F}_{xx})^{-1}.$$

The immediately striking aspect of each of these plots is the minuscule expected error in $\log(M_{\rm B})$, $\log(M_{\rm R})$ and $\log(M_{\rm R0})$, particularly in the right panel. This is to be expected, however, since these are the fractional errors in logarithmic quantities. Translated to errors on the actual mass scales, these correspond to roughly 1% marginalized errors for both selection functions in the absence of selection function error. Additionally, we can see that, while the fractional errors are less impressive for the magnitude-based selection function, the errors using this selection function are more robust against errors in the selection function. In general, we can see that there is not an enormous difference in expected fractional error between the two selection functions for most parameters. The magnitude-based selection function appears to be marginally more robust against increasing selection function error. This probably does not compensate for the short-comings of making the linear assumption shown in Figure 7, but is worth bearing in mind if there is significant error in the photometric redshift (e.g. large errors due to color degeneracies).

For the general halo model parameters, the expected errors are quite small. Indeed, given the disparity between values of α for the NFW and Moore profiles, as well as the scatter on the concentration parameter listed in Bullock et al. (2001), it is possible that the expected errors on these parameters could be smaller than the associated errors from simulations. For the sub-population parameters, the expected errors on η and μ are larger than the Poisson errors on these parameters from the GIF simulations. However, they should be sufficient to determine that they are different from unity, and hence the usefulness of the formalism developed in §2.2. The shape of the Gaussian component (as given by A and A_0) in $\langle N\rangle_{\rm B}$ is not as well determined by the angular correlations as by the simulations, but the errors on the mass scales and power law indices ($M_{\rm R}$, $M_{\rm B}$, $M_{\rm R0}$, $M_{\rm Bs}$, $\gamma_{\rm R}$ and $\gamma_{\rm B}$) should be. Finally, the mass scale for the deviation from a Poisson distribution in the $\langle N(N-1)\rangle$ relation (M_α) should be constrained as well from the angular correlations as in the simulations, but the rate of that deviation (σ_α) is not. As in the case of A and A_0, this constraint may improve given a different set of $\langle N\rangle$ parameters.

7.2. *Parameter Degeneracies*

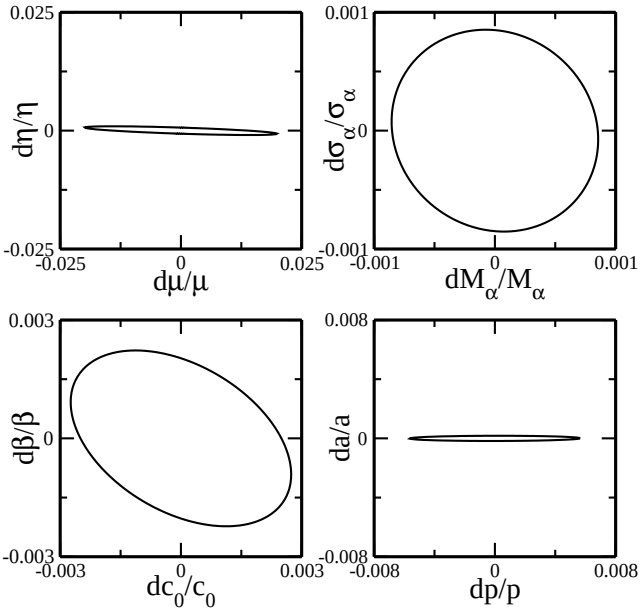

FIG. 13.— Expected error ellipses for four combinations of parameters, normalized by the respective parameter values, for the magnitude-based selection function. Clockwise from upper left: μ vs. η, M_α vs. σ_α, a vs. p and c_0 vs. β. Similar results were obtained for the photometric redshift selection function, with the exception of the μ vs η plot, which was nearly orthogonal.

There are two means by which we can examine the degeneracies between the various halo model parameters. First, we can look at the error ellipses between various parameters we expect to be correlated (e.g. μ & η, c_0 & β, M_α & σ_α). While useful for considering particular pairs of parameters, this approach does not reveal the full extent of the correlations between all the parameters. To examine this, we can decompose the Fisher matrix into its eigenvectors. Provided the Fisher matrix is not singular, these eigenvectors define a basis of orthogonal combinations of parameters and tell us which combinations of halo parameters are naturally constrained by the angular correlations.

The error ellipse for each pair of parameters can be constructed taking the four corresponding elements of **F**, inverting and decomposing the resulting matrix into its eigencomponents. This effectively fixes all of the other parameters in **F**, resulting in errors on each parameter on order those found by taking $(\mathbf{F}_{xx})^{-1/2}$. Figure 13 gives the expected error ellipses (normalized by the respective parameter values) for the three combinations of parameters listed above, as well as the expected error ellipse for a & p, in the limit of a perfect magnitude-based selection function. In all four cases, as with the other 149 combinations of parameters, the shape of the error ellipse is not discernibly different for the two selection functions, although the size will vary between selection functions according to the relative values of δx for the two selection functions. As we can see, our model breaks most of the degeneracy between μ and η seen in the model used in S02, resulting in a nearly orthogonal error ellipse. In contrast, we can see rather strong correlations between the expected errors for our concentration relation.

In addition to looking at correlations between parameters, we can also look at how those correlations vary as we increase the contribution to our error calculations from the selection function. For the case of M_α & σ_α, we can see from Figure 14 that, along with the expected increase in the size of the error ellipse, there is also some degree of wavering in the degree of correlation in the various error regimes. In general, we find that

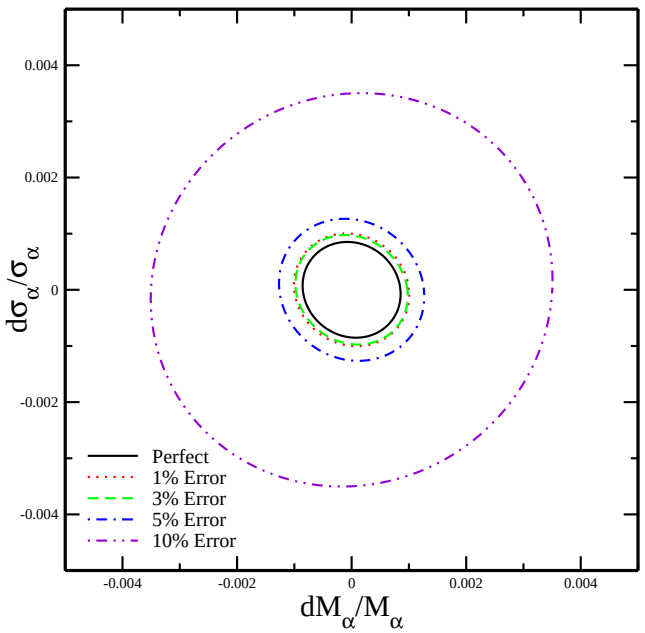

FIG. 14.— Expected error ellipses for M_α & σ_α using the magnitude-based selection function in each of the five selection function error regimes.

this behavior is fairly consistent for all of the various parameter combinations, with the exception of the 10% error regime. In this case, where the covariance on the angular bias is most strongly influenced by the selection function error, there are a number of parameter combinations ($M_{\rm R}$ and $M_{\rm B}$, for instance) where the 10% error ellipses were significantly rotated from the other ellipses, as much as 90 degrees in some cases. Likewise, the behavior of the 10% ellipse for a given pair of parameters for the magnitude-based selection function appeared to be of little use in predicting the orientation of the same error regime in the photometric redshift selection function.

Having figuratively tested the waters with the error ellipses, we can move on to the more daunting task of examining the eigenbases for the various Fisher matrices to determine the parameter combinations that our measurements best constrain. Tables 3 and 4 give the eigenbases for the two selection functions in the limit of no selection function error. In both cases, we have ignored those parameters in each eigenmode which contribute less than 2% to the total amplitude of the mode (i.e. all those parameters whose eigenmode coefficient e_i was $|e_i| \leq 0.14$). As one might have guessed, there are relatively few modes that are simply determined by one or two halo parameters. The exceptions to this rule occur for the modes with the best and worst constraints, where we find strong constraints on simple combinations of a few parameters ($\gamma_{\rm R}$, a and β) contrasted with relatively weaker constraints on single parameters (c_0, η, A_0 and μ). Not surprisingly, members of this second set tend to also have the worst fractional unmarginalized errors and vice versa for the elements of the first set.

Likewise, we can see some reflection of the error ellipses in the eigenbasis. The parameter pairs μ & η and a & p were very nearly uncorrelated in Figure 13 and we can see that these combinations of parameters do not appear in any of the eigenmodes for either of the Fisher matrices. Conversely, the M_α & σ_α pair almost invariably appear together in the eigenmodes and with the expected relative signs and amplitudes. This is not a perfect guide, however, given the correlation between c_0 and β in the error ellipses and the absence of an eigenmode containing both parameters.

When we add in errors due to selection function uncertainty, the effects are similar to those seen in the error ellipses. The eigenmodes remain mostly unchanged with increasing selection function error; the contribution of each parameter to a given eigenmode remains identical within $\sim 2\%$ of the total eigenmode amplitude and appears to preserve the relative signs. Likewise, we see an increase in the error on each eigenmode as the selection function error increases; errors increase by a factor of 2 for each 5% of selection function error. The one case of significant change in the eigenbasis we do see is in the best constrained eigenmodes; the degeneracy between $\gamma_{\rm R}$, a and β changes for both selection functions, leading to an idependent mode for a in the magnitude-base selection function and independent modes for both $\gamma_{\rm R}$ and a in the photometric case.

8. CONCLUSION

As hoped, the eventual angular clustering measurements for red and blue galaxies should provide strong constraints ($\sim 1-10\%$) on a wide variety of halo model parameters. This remains the case for moderate levels of uncertainty in the selection function. In the limit of small selection function errors, we can achieve similar constraints using both a magnitude and photometric redshift based selection functions. This second approach

also allows us to have much greater confidence that we are constraining a simple $\langle N \rangle$ relation, rather than a weighted projection over the $\langle N \rangle$ surfaces in Figure 2. In either case, we can see that the parameters relating to the shape of the Gaussian part of $\langle N \rangle_{\rm B}$ do not strongly affect the angular clustering, but the mass scales and power law indices in $\langle N \rangle_{\rm R}$ & $\langle N \rangle_{\rm B}$ should be very tightly constrained. In addition, we should have sufficient constraints on the parameters relating the halo mass function, profile, and concentration to determine if the values measured from simulations are consistent with observations. Finally, the completed SDSS measurements should be sufficient to determine if the method for distributing red and blue galaxies within a halo given in S02 accurately reproduces the observed clustering.

In addition to the angular correlation function work presented here, there are a number of other measurements one might adopt to constrain halo parameters in a similar fashion to the methodology presented here. In particular, measurements of both strong and weak lensing should provide a great deal of information about the structure of dark matter halos and the subsidiary information in rich galaxy surveys should be sufficient to probe different galaxy biasing by type and color on physical scales, rather than the angular ones presented here. Likewise, one could readily construct alternative formulations of the power spectrum (weighting by star formation rate or bulge-to-disk ratios, for instance) which could be used in conjunction with spectroscopic and morphological information to inform future simulations. Deeper spectroscopic surveys could also attempt to find the break in the galaxies power spectrum shown in Figure 4 as the fraction of galaxies shifts toward the blue at higher redshifts. Finally, one could also consider the clustering of galaxy clusters as a separate test of the linear behavior of the halo concentration and mass function.

We would like to thank Scott Dodelson, Josh Frieman, Wayne Hu and Stephen Kent for many useful suggestions regarding the text and scope of this paper. Likewise, we would like to thank Ravi Sheth, Bhuvnesh Jain and Roman Scoccimarro for a number of very useful conversations. Additional thanks to Guinevere Kauffmann, Volker Springel and Antonaldo Diaferio for assistance with the simulation data.

Support for this work was provided by the NSF through grant PHY-0079251 as well as by NASA through grant NAG 5-7092 and the DOE.

9. REFERENCES

Benson, A.J., Cole, S., Frenk, C.S., Baugh, C.M., & Lacey, C.G., 2000, mnras, 311, 793
Budavári, T. ;., Szalay, A. S., Connolly, A. J., Csabai, I. ;., & Dickinson, M. 2000, AJ, 120, 1588
Connolly, A. et al., 2001, submitted to ApJ (astro-ph/0107417)
Cooray, A. & Hu, W., 2001, ApJ , 554, 56
Couchman, H.M.P, Thomas, P.A, & Pearce, F.R., 1995, ApJ , 452, 797
Diaferio, A., Kauffmann, G., Colberg, J.M., & White, S.D.M., 1999, MNRAS , 307, 537
Dodelson, S. et al., 2001, submitted to ApJ (astro-ph/0107421)
Eisenstein, D.J. & Zaldarriaga, M. 2001, ApJ , 546, 2
Fukugita, M., Ichikawa, T., Gunn, J.E., Doi, M., Shimasaku, K., & Scheider, D.P., 1996, AJ, 111,1748
Gaztanaga, E., 2001, astro-ph/0106379
Gunn, J. E., & The SDSS Collaboration, 1998,AJ, 116, 3040
Kauffmann, G., Colberg, J.M., Diaferio, A. & White, S.D.M., 1999, MNRAS , 303, 188
Lin, H., Yee, H. K. C., Carlberg, R. G., Morris, S. L., Sawicki, M., Patton, D. R., Wirth, G., & Shepherd, C. W., 1999, ApJ , 518, 533
Jing, Y.P., Mo, H.J., Borner, G., 1998, ApJ , 499,20
Ma, C.-P. & Fry, J.N., 2000, ApJ , 543, 503
Mo, H.J., White, S.D.M., 1996, MNRAS , 282, 1096
Moore, B., Governato, F., Quinn, T., Stadel, J., & Lake, G., 1999, MNRAS , 261, 827
Navarro, J., Frenk, C., & White, S.D.M., 1996, ApJ , 462, 563
Navarro, J., Frenk, C., & White, S.D.M., 1997, ApJ , 490, 493
Peacock, J.A. & Smith, R.E., 2000, MNRAS , 318, 1144
Press, W.H. & Schechter, P., 1974, ApJ , 187, 425
Pryke, C., Halverson, N.W., Leitch, E.M., Kovac, J., Carlstrom, J.E., Holzapfel, W.L., & Dragovan, M. 2002, ApJ , 568, 46.
Scoccimarro, R., Zaldarriaga, M., & Hui, L., 1999, ApJ , 527, 1
Scoccimarro, R., Sheth, R.K., Hui, L. & Jain, B., 2001, ApJ , 546, 20
Scranton, R., 2001, submitted to MNRAS , (astro-ph/0108266)
Scranton, R. et al., 2002, submitted to ApJ (astro-ph/0107416)
Seljak, U., 2000, MNRAS , 318, 203
Sheth, R. K. & Jain, B. 1997, MNRAS , 285, 231
Sheth, R., Hui, L., Diaferio, A., Scoccimarro, R., 2001, MNRAS , 325, 1288
Sheth, R. & Tormen, G., 1999, MNRAS , 308, 119
Tegmark, M., Hamilton, A.J.S., Xu, Y. 2001, submitted to MNRAS , (astro-ph/0111575)
White, M., Hernquist, L., Springel, V., 2001, ApJ , 550, 129
White, S.D.M & Rees, M.J., 1978, MNRAS , 183, 341
York, D. G., & The SDSS Collaboration 2000, AJ120, 157

TABLE 3
FISHER EIGENMODES FOR MAGNITUDE-BASED SELECTION FUNCTION.

Parameter Basis	Expected Error
$0.32\gamma_R + 0.84a + 0.43\beta$	$\pm 1.13 \times 10^{-4}$
$0.55\gamma_R - 0.53a + 0.62\beta$	$\pm 1.85 \times 10^{-4}$
$0.77\gamma_R - 0.63\beta$	$\pm 6.79 \times 10^{-4}$
$-0.37\log(M_B) - 0.19\log(M_{R0}) + 0.53\gamma_B - 0.28A + 0.55M_{Bs} + 0.27\alpha + 0.16\beta$	$\pm 8.99 \times 10^{-4}$
$0.24\log(M_R) + 0.21\gamma_B - 0.19A + 0.90M_\alpha$	$\pm 2.29 \times 10^{-3}$
$0.46\log(M_R) + 0.39\log(M_{R0}) - 0.30\gamma_B + 0.24M_{Bs} + 0.48\sigma_\alpha + 0.40\alpha - 0.26a$	$\pm 2.88 \times 10^{-3}$
$-0.75\log(M_{R0}) - 0.27\gamma_B + 0.24M_{Bs} + 0.18M_\alpha + 0.36\sigma_\alpha + 0.28p$	$\pm 3.75 \times 10^{-3}$
$-0.67\log(M_R) - 0.18\gamma_B + 0.23M_{Bs} + 0.33M_\alpha + 0.27\sigma_\alpha - 0.27\alpha - 0.43p$	$\pm 4.37 \times 10^{-3}$
$0.34\log(M_R) + 0.31\log(M_{R0}) + 0.19\gamma_B + 0.38M_{Bs} - 0.73\alpha + 0.21p$	$\pm 4.92 \times 10^{-3}$
$-0.32\log(M_R) + 0.33\log(M_{R0}) + 0.20\gamma_B + 0.29A + 0.37\sigma_\alpha + 0.66p$	$\pm 5.82 \times 10^{-3}$
$0.41\log(M_B) + 0.62\gamma_B + 0.29A - 0.33M_{Bs} + 0.33\sigma_\alpha - 0.32p$	$\pm 8.21 \times 10^{-3}$
$0.48\log(M_B) + 0.23M_{Bs} - 0.19\sigma_\alpha + 0.79c_0$	$\pm 1.66 \times 10^{-2}$
$-0.55\log(M_B) - 0.36M_{Bs} + 0.31\sigma_\alpha - 0.31\alpha - 0.20p + 0.52c_0$	$\pm 2.04 \times 10^{-2}$
$-0.31\log(M_B) + 0.80A + 0.26M_{Bs} - 0.35\sigma_\alpha - 0.18p$	$\pm 3.37 \times 10^{-2}$
$0.97\eta + 0.20c_0$	± 0.11
A_0	± 0.61
μ	± 1.82

TABLE 4
FISHER EIGENMODES FOR PHOTOMETRIC REDSHIFT SELECTION FUNCTION.

Parameter Basis	Expected Error
$0.85\gamma_R - 0.52a$	$\pm 1.44 \times 10^{-4}$
$0.52\gamma_R + 0.85a$	$\pm 2.76 \times 10^{-4}$
β	$\pm 4.89 \times 10^{-4}$
$-0.29\log(M_B) - 0.17\log(M_{R0}) + 0.40\gamma_B - 0.18A + 0.35M_{Bs} + 0.20\sigma_\alpha + 0.71\alpha$	$\pm 8.23 \times 10^{-4}$
$-0.30\log(M_B) - 0.18\log(M_{R0}) + 0.42\gamma_B - 0.24A + 0.41M_{Bs} - 0.66\alpha$	$\pm 1.15 \times 10^{-3}$
$0.28\log(M_{R0}) - 0.14\gamma_B + 0.26M_{Bs} - 0.21M_\alpha + 0.83\sigma_\alpha - 0.17\alpha - 0.15p$	$\pm 2.28 \times 10^{-3}$
$0.16\log(M_R) + 0.25\gamma_B - 0.29M_{Bs} + 0.80M_\alpha + 0.37\sigma_\alpha$	$\pm 3.38 \times 10^{-3}$
$-0.15\log(M_B) + 0.68\log(M_R) - 0.42\log(M_{R0}) - 0.30\gamma_B + 0.23M_{Bs} + 0.42p$	$\pm 3.91 \times 10^{-3}$
$0.25\log(M_B) + 0.18\log(M_R) - 0.34\gamma_B - 0.39A + 0.46M_{Bs} + 0.41M_\alpha - 0.17\sigma_\alpha - 0.46p$	$\pm 4.23 \times 10^{-3}$
$0.56\log(M_B) - 0.19\log(M_R) + 0.19\log(M_{R0}) + 0.16\gamma_B - 0.22A + 0.27M_{Bs} + 0.68p$	$\pm 5.19 \times 10^{-3}$
$0.32\log(M_R) + 0.50\log(M_{R0}) + 0.25\gamma_B + 0.59A + 0.33M_{Bs} + 0.18M_\alpha - 0.24\sigma_\alpha$	$\pm 7.54 \times 10^{-3}$
$-0.45\log(M_R) - 0.49\log(M_{R0}) - 0.22\gamma_B + 0.53A + 0.31M_{Bs} + 0.29M_\alpha$	$\pm 8.87 \times 10^{-3}$
$0.60\log(M_B) + 0.33\log(M_R) - 0.36\log(M_{R0}) + 0.48\gamma_B + 0.18A - 0.32p$	$\pm 1.42 \times 10^{-2}$
c_0	$\pm 2.76 \times 10^{-2}$
η	± 0.16
A_0	± 0.34
μ	± 3.18

APPENDIX A

HALO MODEL POWER SPECTRA

Rewriting Equation 1 in terms of the concentration and the mass, we get

$$\rho(r,M) = \frac{\rho_S}{(rc/r_v)^{-\alpha}(1+rc/r_v)^{3+\alpha}}, \quad \text{(A1)}$$

where

$$r_v^3 = \frac{3M}{4\pi\Delta_V\bar{\rho}}, \quad \text{(A2)}$$

$$\rho_S = \frac{\Delta_V\bar{\rho}c^3(M)}{3}\left[\int_0^{c(M)} d\chi \frac{\chi^{2+\alpha}}{(1+\chi)^{3+\alpha}}\right]^{-1}, \quad \text{(A3)}$$

$\bar{\rho}$ is the mean matter density and M is the mass of the halo. Since we will be working in wavenumber space when we generate the power spectrum, we actually need to consider the Fourier transform of the halo profile,

$$y(k,M) = \frac{1}{M}\int_0^{r_v} 4\pi r^2 \rho(r,M)\frac{\sin(kr)}{kr}dr, \quad \text{(A4)}$$

where we have normalized over mass so that $y(0,M) = 1$ and $y(k>0,M) < 1$. Note that this implies that $\rho(r > r_v) = 0$, truncating the mass integration at the virial radius. This condition can be relaxed, provided that one scales the halo mass appropriately.

With this in hand, we can move on to the next component of the halo model, the halo mass function, (dn/dM). The form of the mass function is given in Equation 7, but we need to properly normalize it by requiring that

$$\frac{1}{\bar{\rho}}\int_0^\infty \frac{dn}{dM}MdM = \int f(\nu)d\nu = 1, \quad \text{(A5)}$$

for the dark matter distribution. On nonlinear scales, we expect the halos to cluster more strongly than the mass, and vice versa for linear scales (Mo & White, 1996). This means we need to positively bias the clustering of the high mass halos relative to the low mass halos. We can generate this sort of halo biasing scheme for the ST mass function using

$$b(\nu) = 1 + \frac{\nu' - 1}{\delta_c} + \frac{2p}{\delta_c(1+\nu'^p)}. \quad \text{(A6)}$$

In order for the eventual power spectrum to reduce to a linear power spectrum on large scales, we need to impose the further constraint that

$$\int_0^\infty f(\nu)b(\nu)d\nu = 1, \quad \text{(A7)}$$

requiring that the biased halos with mass greater than M_* be balanced out by anti-biased halos with mass less than M_*. This integral is satisfied automatically if we use Equation A6 and have properly normalized $f(\nu)$.

Using just these three components, we can generate the power spectrum for the dark matter. However, in order to predict the galaxy power spectrum, we need to know how many galaxies are in a given halo (under the assumption that the distribution of galaxies in the halo follows the halo profile). These $\langle N\rangle$ relations are given in Equations 8 and 9. The inclusion of galaxies does change the normalization of Equation A5 to

$$\int_0^\infty \frac{\langle N\rangle}{M(\nu)}f(\nu)d\nu = \frac{\bar{n}}{\bar{\rho}}, \quad \text{(A8)}$$

where $\bar{n}$ is the mean number of galaxies and $\bar{\rho}$ is the mean matter density at redshift z.

On large scales, the power spectrum is dominated by correlations between galaxies in separate halos. We need to convolve the halo profile with the mass function to account for the fact that halos are not point-like objects. Since we are in Fourier space, we can perform the convolution using simple multiplication. The halo-halo power ($P_{GG}^{hh}(k)$) is then simply,

$$P_{GG}^{hh}(k) = P_{LIN}(k)\left[\frac{\bar{\rho}}{\bar{n}}\int_0^\infty f(\nu)\frac{\langle N\rangle}{M(\nu)}b(\nu)y(k,M)d\nu\right]^2, \quad \text{(A9)}$$

where $P_{LIN}(k)$ is the linear dark matter power spectrum,

$$\left\langle \delta(\mathbf{k})\delta(\mathbf{k}')\right\rangle = (2\pi)^3\delta\left(\mathbf{k}-\mathbf{k}'\right)P_{LIN}(k). \quad \text{(A10)}$$

For small scales, the dominant contribution to the power spectrum comes from correlations between galaxies within the same halo. This single halo term is independent of k at larger scales, giving it a Poisson-like behavior. In order to account for the fact that a single galaxy within a halo does not correlate with itself, we use the second moment of the galaxy number relation, $\langle N(N-1)\rangle$, to calculate the Poisson power ($P_{GG}^{P}(k)$),

$$P_{GG}^{P}(k) = \frac{\bar{\rho}}{\bar{n}^2}\int_0^\infty f(\nu)\frac{\langle N(N-1)\rangle}{M}|y(k,M)|^\zeta d\nu. \quad \text{(A11)}$$

Seljak (2000) takes $\zeta = 2$ for $\langle N(N-1)\rangle > 1$ and $\zeta = 1$ for $\langle N(N-1)\rangle < 1$; this is done to account for the galaxy at the center of the halo in the limit of small number of galaxies. Adding $P_{GG}^{hh}(k)$ and $P_{GG}^{P}(k)$, we recover the galaxy power spectrum at all wavenumbers, $P_{GG}(k)$.

A.1. *Calculating Subpopulation Power Spectra*

With the modifications to the mass distributions which go into red and blue galaxies given by Equations 13 and 14, we need to regenerate $y(k,M)$ for each profile. We also need to renormalize $f(\nu)$ for each sub-population using Equation A8 to account for the differences in $\langle N\rangle$ and $\bar{n}$:

$$\int_0^\infty \frac{\langle N\rangle_R}{M(\nu)}f_R(\nu)d\nu = \frac{\bar{n}_R}{\bar{\rho}} \quad \text{(A12)}$$
$$\int_0^\infty \frac{\langle N\rangle_B}{M(\nu)}f_B(\nu)d\nu = \frac{\bar{n}_B}{\bar{\rho}}$$

Once this has been done, we can insert the above (along with the color-dependent halo profiles) into Equations A9 and A11 to generate the power spectra for red and blue galaxies:

$$\frac{P_{RR}^{hh}(k)}{P_{LIN}(k)} = \left[\frac{\bar{\rho}}{\bar{n}_R}\int_0^\infty f_R(\nu)\frac{\langle N\rangle_R}{M(\nu)}b(\nu)y_R(k,M)d\nu\right]^2 \quad \text{(A13)}$$
$$P_{RR}^{P}(k) = \bar{\rho}\int_0^\infty f_R(\nu)\frac{\langle N(N-1)\rangle_R}{\bar{n}_R^2 M(\nu)}|y_R(k,M)|^\zeta d\nu,$$
$$\frac{P_{BB}^{hh}(k)}{P_{LIN}(k)} = \left[\frac{\bar{\rho}}{\bar{n}_B}\int_0^\infty f_B(\nu)\frac{\langle N\rangle_B}{M(\nu)}b(\nu)y_B(k,M)d\nu\right]^2$$
$$P_{BB}^{P}(k) = \bar{\rho}\int_0^\infty f_B(\nu)\frac{\langle N(N-1)\rangle_B}{\bar{n}_B^2 M(\nu)}|y_B(k,M)|^\zeta d\nu.$$

As before, we generate the total power spectra ($P_{RR}(k)$ and $P_{BB}(k)$) by taking the sum of these parts,

$$P_{RR}(k) = P_{RR}^{hh}(k) + P_{RR}^{P}(k) \quad \text{(A14)}$$
$$P_{BB}(k) = P_{BB}^{hh}(k) + P_{BB}^{P}(k).$$

APPENDIX B

SUB-POPULATION SELECTION AND $\langle N \rangle$ PARAMETERS

For all of the parameter values related to the red and blue galaxy HOD, we fit the relations in Equations 8, 9, and 11 to the galaxy catalogs produced by the GIF simulations. The details of the semi-analytic methods applied in the simulations can be found in Kaufmann et al. (1999), but we will briefly discuss some of the relevant features here.

In the broadest strokes, semi-analytic methods like those applied in the GIF simulations take the outputs of an N-body cosmological simulation at a number of time-steps, determine where galaxies will have formed based on some prescription and let the galaxies evolve from that point in time until the present. In the case of the GIF simulations, the N-body simulations were generated using *Hydra* (Couchman, Thomas & Pearce, 1995), an adaptive particle-particle particle mesh code written as part of the VIRGO collaboration. Four different cosmological models were used in the initial work, but, for the purposes of the calculations in this paper, we only used the ΛCDM outputs ($\Omega_M = 0.3, \Omega_\Lambda = 0.7, h = 0.7, \sigma_8 = 0.9$) with the SDSS filters (not mentioned in Kaufmann et al.). These simulations are 141 h^{-1} Mpc on a side and have a mass resolution on order $10^{11} h^{-1} M_\odot$.

At each time step used in the GIF simulations, a friends-of-friends group finding routine with linking length 0.2 was applied to the N-body outputs and each group of 10 or more particles was marked as a dark matter halo. The most bound member of each such group was flagged as the central galaxy of the halo. In subsequent time steps, previously determined halos are checked against mergers with other halos. In the case of a merger, the galaxy at the center of the most massive progenitor halo is considered to be at the center of the combined halo and the properties of that galaxy are transferred to the most bound particle of the new halo. Galaxies associated with less massive progenitor halos are now satellite galaxies of the new halo and remain associated with their original particles. In the GIF simulation outputs used for our calculations, each catalog contained $\sim 90,000$ halos and $\sim 180,000$ galaxies (of which $\sim 35,000$ fell within our apparent magnitude cut).

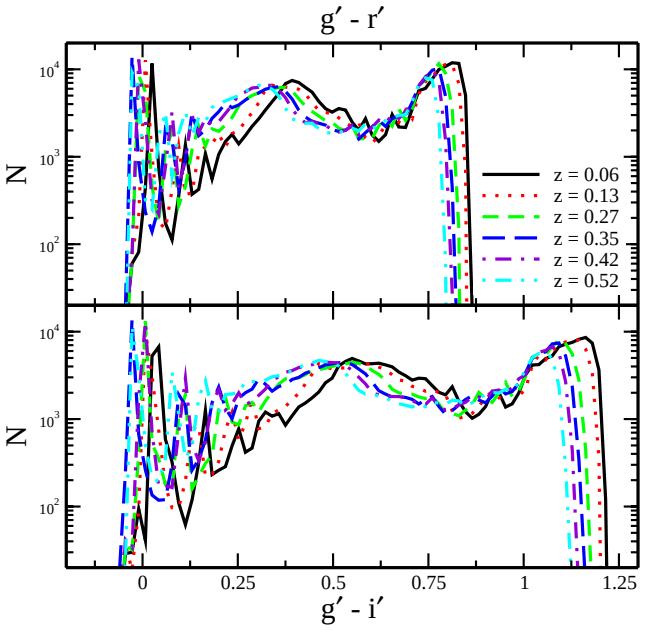

FIG. 15.— Color distributions as a function of redshift using $g' - r'$ (upper) and $g' - i'$ (lower).

Once the positions of the individual galaxies within each halo have been set, the evolution of the stars in each galaxy can be determined. Even in the simplest terms, this requires a number of considerations: availability of cool gas, star formation rate, supernovae feedback, initial stellar mass function, metallicity (this is held at solar levels throughout the calculations), etc. In addition to these intra-galactic effects, there are also merger effects (combination of two satellite galaxies or in-fall of satellite galaxies into the central halo galaxy) and the associated creation of galaxy bulges and star-burst activity. All of these processes require tuning to one degree or another in order to reasonably reproduce observed luminosity functions and Tully-Fisher relations. Since future SDSS angular clustering measurements discussed here will combine both the galaxy evolution and clustering aspects of the model (at a variety of redshifts), they should serve as an excellent test for many aspects of these treatments.

In splitting the GIF simulation galaxy catalogs into red and blue samples, we had two primary considerations. First, we wanted to produce a selection method for the data that was robust in segregating what appear to be two rather distinct subpopulations. Second, we wanted a criterion which could reasonably be applied to actual galaxies near the limit of our magnitude selection of $r' = 21$. The first of these requirements meant choosing a color cut that varied between the two populations slowly enough that passive evolution in galaxy colors over the extent of the redshift range was relatively small. The other requirement meant restricting ourselves to the g', r' and i' bands, as objects at the faint end of our magnitude cut will often fall below the detection threshold in u' and z'.

The data sets we considered for this selection consisted of 6 redshift epochs: $z = 0.06$, 0.13, 0.27, 0.35, 0.42 and 0.52. Figure 15 shows the distribution of rest-frame $g' - r'$ and $g' - i'$ colors for the galaxies at each of the redshift epochs. In both cases, the distribution is roughly bimodal, with a spike of very blue star-forming galaxies at $g' - r' = g' - i' \sim 0$. There is not an enormous difference between the two color distributions and it is clear that a simple straight line cut will select a slightly different population at higher redshift than at lower redshift. However, given the wider distribution of $g' - i'$, we should suffer from less difference with redshift than with $g' - r'$. With this in mind, we split our sample at

$$g' - i' = 0.85. \tag{B1}$$

We can test the effectiveness of our color selection by looking at the variation of the $\langle N \rangle$ surfaces as a function of redshift. In Figure 16, we plot the surfaces for the whole galaxy sample. As mentioned in §2, the overall galaxy surface appears reasonably static as a function of redshift and the sub-population surfaces behave similarly. We can also see that the shifting in apparent magnitude proceeds with redshift as we would expect.

For a more specific look at the possible evolution of the $\langle N \rangle$ relations as it applies to the calculations in this paper, we can shift the surface from each redshift regime appropriately for the $z \sim 0.3$ selection function, apply our magnitude cut and compare the $\langle N \rangle$ curves. Figure 17 shows the $\langle N \rangle$ relations produced by this method for the $z = 0.06$, 0.13, 0.27 and 0.35 epochs. The $\langle N \rangle_B$ curves show no significant signs of evo-

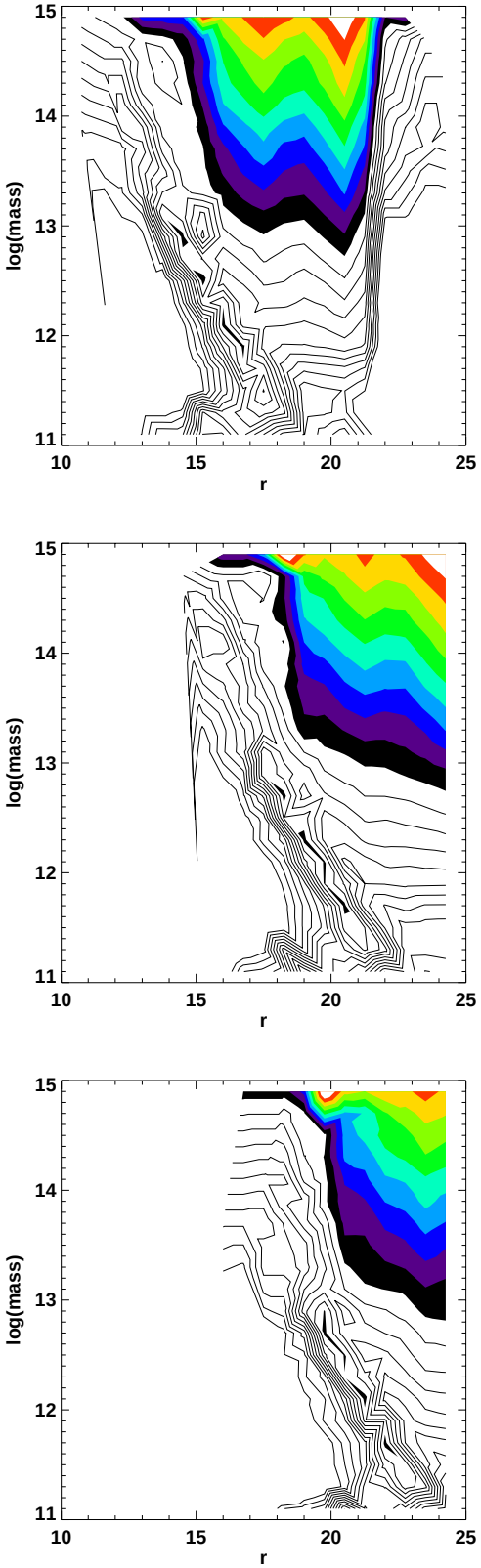

FIG. 16.— $\langle N \rangle$ as a function of halo virial mass and r' magnitude from the GIF simulations at $z = 0.06$, 0.27 and 0.52 (top to bottom). As in Figure 2, filled contours indicate $\log(\langle N \rangle) > 0$, wire-frame contours indicate $\log(\langle N \rangle) < 0$ and successive contours a change of 0.25 in $\log(\langle N \rangle)$.

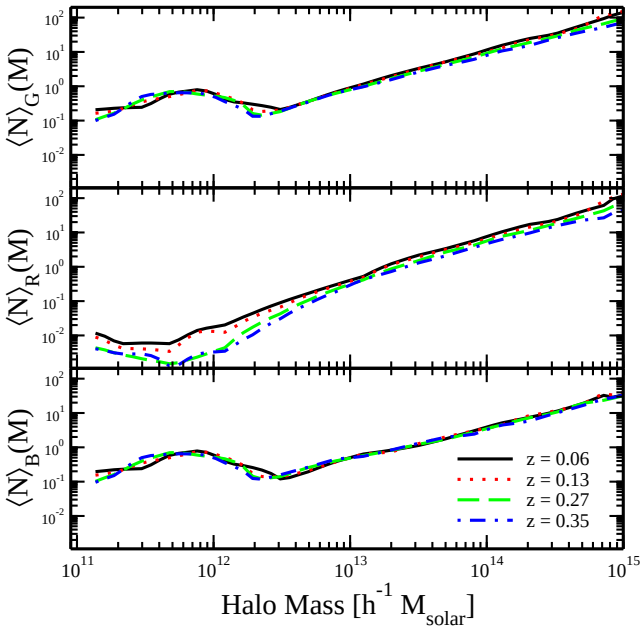

FIG. 17.— $\langle N \rangle$ relations for all galaxies (upper), red galaxies (middle) and blue galaxies (bottom).

lution. There is some shift in the mass scales for $\langle N \rangle_R$, both M_R and M_{R0}, but the γ_R is largely unchanged. In choosing our fiducial model, we fit the $\langle N \rangle_R$ parameters from the $z = 0.27$ epoch. While this behavior does make the modeling of the $\langle N \rangle$ relations more complicated than would be the case if the all the surfaces were static, the prediction of a red galaxy distribution evolving against a background of a static blue galaxy distribution is intriguing.

In addition to the $\langle N \rangle$ relations, we can also check the $\langle N(N-1) \rangle$ relations for evolution. As one might expect from Figure 17, measurements of $\langle N(N-1) \rangle_R$ from the simulations were very noisy and did not lend themselves to a reliable fit in the region where one expects strong deviation from a Poisson distribution. Rather, since we chose to use a universal α_M, we fit M_α and σ_α from Equation 11 using $\langle N(N-1) \rangle$ and $\langle N \rangle$ measurement of the blue sub-sample. This approach may miss some of the evolution present in the red galaxy sample, but this seems unavoidable. Unlike in the $\langle N \rangle$ measurements, more recent redshift epochs did show stronger Poisson behavior at lower mass scales; M_α stayed roughly fixed, but σ_α increased, leading to a more gradual decrease in α_M for lower mass halos. As with $\langle N \rangle_R$, the effect was not dramatic, but did result in roughly doubling σ_α from the $z = 0.06$ epoch from its value for the $z = 0.35$ epoch. As before, we chose the parameter fits using the $z = 0.27$ surfaces for our fiducial model.