# The FMOS-COSMOS Survey of Star-forming Galaxies at $Z \sim 1.6$. V: Properties of Dark Matter Halos Containing Hα Emitting Galaxies

Daichi Kashino[1,2], Surhud More[3], John D. Silverman[3], Emanuele Daddi[4], Alvio Renzini[5], David B. Sanders[6], Giulia Rodighiero[7], Annagrazia Puglisi[7,8], Masaru Kajisawa[9,10], Francesco Valentino[4,11], Jeyhan S. Kartaltepe[12], Olivier Le Fèvre[13], Tohru Nagao[9], Nobuo Arimoto[14,15], and Naoshi Sugiyama[2,3]

[1] Institute for Astronomy, Department of Physics, ETH Zürich, Wolfgang-Pauli-Strasse 27, CH-8093 Zürich, Switzerland; kashinod@phys.ethz.ch
[2] Division of Particle and Astrophysical Science, Graduate School of Science, Nagoya University, Furo-cho, Chikusa-ku, Nagoya, Aichi 464-8602, Japan
[3] Kavli Institute for the Physics and Mathematics of the Universe (WPI), Todai Institutes for Advanced Study,
the University of Tokyo, Kashiwanoha, Kashiwa, Chiba 277-8583, Japan
[4] Laboratoire AIM-Paris-Saclay, CEA/DSM-CNRS-Université Paris Diderot, Irfu/Service d'Astrophysique, CEA-Saclay,
Orme des Merisiers, F-91191 Gif-sur-Yvette, France
[5] INAF Osservatorio Astronomico di Padova, vicolo dell'Osservatorio 5, I-35122 Padova, Italy
[6] Institute for Astronomy, University of Hawaii, 2680 Woodlawn Drive, Honolulu, HI 96822, USA
[7] Dipartimento di Fisica e Astronomia, Università di Padova, vicolo dell'Osservatorio, 2, I-35122 Padova, Italy
[8] ESO, Karl-Schwarzschild-Straße 2, D-85748 Garching bei München, Germany
[9] Research Center for Space and Cosmic Evolution, Ehime University, Bunkyo-cho 2-5, Matsuyama, Ehime 790-8577, Japan
[10] Graduate School of Science and Engineering, Ehime University, Bunkyo-cho 2-5, Matsuyama, Ehime 790-8577, Japan
[11] Dark Cosmology Centre, Niels Bohr Institute, University of Copenhagen, Juliane Maries Vej 30, DK-2100 Copenhagen, Denmark
[12] School of Physics and Astronomy, Rochester Institute of Technology, 84 Lomb Memorial Drive, Rochester, NY 14623, USA
[13] Aix-Marseille Université, CNRS, LAM (Laboratoire d'Astrophysique de Marseille), UMR 7326, F-13388 Marseille, France
[14] Astronomy Program, Department of Physics and Astronomy, Seoul National University 599 Gwanak-ro, Gwanak-gu, Seoul, 151-742, Korea
[15] Subaru Telescope, National Astronomical Observatory of Japan, National Institutes of Natural Sciences (NINS), 650 North Aohoku Place, Hilo, HI 96720, USA

## Abstract

We study the properties of dark matter halos that contain star-forming galaxies at $1.43 \leqslant z \leqslant 1.74$, using the FMOS-COSMOS survey. The sample consists of 516 objects with a detection of the Hα emission line, which represent the star forming population at this epoch, having a stellar mass range of $10^{9.57} \leqslant M_*/M_\odot \lesssim 10^{11.4}$ and a star-formation rate range of $15 \lesssim \mathrm{SFR}/(M_\odot \, \mathrm{yr}^{-1}) \lesssim 600$. We measure the projected two-point correlation function while carefully taking into account observational biases, and find a significant clustering amplitude at scales of $0.04$–$10 \, h^{-1}$ cMpc, with a correlation length $r_0 = 5.26^{+0.75}_{-0.62} \, h^{-1}$ cMpc and a bias $b = 2.44^{+0.38}_{-0.32}$. We interpret our clustering measurement using a halo occupation distribution model. The sample galaxies appear to reside in halos with mass $M_{\rm h} = 4.71^{+1.19}_{-1.62} \times 10^{12} \, h^{-1} \, M_\odot$ on average, which will likely become present-day halos of mass $M_{\rm h} \, (z=0) \sim 2 \times 10^{13} \, h^{-1} M_\odot$, equivalent to the typical halo mass scale of galaxy groups. We then confirm the decline of the stellar-to-halo mass ratio at $M_{\rm h} < 10^{12} \, M_\odot$, finding $M_*/M_{\rm h} \approx 5 \times 10^{-3}$ at $M_{\rm h} = 7.5 \times 10^{11} \, M_\odot$, which is lower by a factor of 2–4 than those measured at higher masses ($M_{\rm h} \sim 10^{12-13} \, M_\odot$). Finally, we use our results to illustrate the future capabilities of Subaru's Prime-Focus Spectrograph, a next-generation instrument that will provide strong constraints on the galaxy-formation scenario by obtaining precise measurements of galaxy clustering at $z > 1$.

*Key words:* galaxies: evolution – galaxies: formation – galaxies: halos

## 1. Introduction

According to our current cosmological model (e.g., White & Rees 1978; Springel et al. 2006; Frieman et al. 2008), the formation of structure in the Universe is dictated by cold dark matter and dark energy. On large scales, matter is organized in cosmic web-like structures that consist of dense nodes, filaments, sheets, and voids. Dark matter virializes into extended halos at the density peaks within this structure. Galaxies form within these halos as a result of complex baryonic processes, such as star formation and feedback (Rees & Ostriker 1977; White & Rees 1978). Understanding the connection between galaxies and dark matter halos can thus elucidate the physical processes that lead to the formation and evolution of galaxies. Specifically, the mass of halos sets their global properties, such as the abundance and spatial distribution (Press & Schechter 1974; Cole & Kaiser 1989; Mo & White 2002), as well as the evolution of their constituent galaxies. A number of different methods have been used to infer halo masses and constrain the connection between galaxies and dark matter halo masses. Direct techniques, such as measurements of galaxy rotation curves (e.g., Rubin et al. 1980, 1982) and X-ray emission from the hot intracluster gas (e.g., Rykoff et al. 2008), as well as statistical techniques, such as the kinematics of satellite galaxies (e.g., More et al. 2009, 2011) and weak gravitational lensing (e.g., Hoekstra et al. 2005; Mandelbaum et al. 2006; Cacciato et al. 2009; Leauthaud et al. 2011; Han et al. 2015; Mandelbaum et al. 2016; van Uitert et al. 2016) can be used to evaluate dynamical mass of systems—albeit all restricted to fairly local redshifts. At redshifts greater than unity, one must resort to indirect techniques, such as subhalo abundance matching, to infer the galaxy–dark matter connection (e.g., Kravtsov et al. 2004; Conroy et al. 2006; Kravtsov et al. 2006; Behroozi et al. 2010; Moster et al. 2010; Masaki et al. 2013).

Analyses of spatial distribution of galaxies can also help to constrain the galaxy–dark matter connection, as galaxies reflect the spatial clustering properties of the halos. For a given cosmological model, the clustering of dark matter halos

depends upon halo mass, such that more-massive halos display stronger clustering than less-massive halos. Therefore, the halo masses can be inferred from the observed clustering amplitude of the hosted galaxies. Large galaxy surveys, such as the two-degree-Field Galaxy Redshift Survey (Colless et al. 2001; Hawkins et al. 2003) and Sloan Digital Sky Survey (SDSS; York et al. 2000), have successfully enabled studies of galaxy clustering at low redshift ($z \sim 0.1$; e.g., Park et al. 1994; Guzzo et al. 2000; Norberg et al. 2001; Zehavi et al. 2002). These studies have also shown that galaxy clustering depends on galactic properties, such as luminosity, color, and morphology, such that relatively luminous (i.e., massive), redder, and bulge-dominated galaxies cluster more strongly (or live in more massive halos), whereas less-massive, bluer, or disk-dominated galaxies have a weaker clustering signal (e.g., Norberg et al. 2002; Zehavi et al. 2005, 2011). Beyond the local Universe ($z \gtrsim 0.5$), most studies of galaxy clustering have measured angular correlation functions based on photometric redshift or color selection (e.g., Brown et al. 2003; Ouchi et al. 2004; Adelberger et al. 2005; Hartley et al. 2008; McCracken et al. 2010; Wake et al. 2011; Coupon et al. 2012; McCracken et al. 2015), whereas measurements based on spectroscopic samples have been limited (e.g., Coil et al. 2004; Meneux et al. 2008, 2009; Abbas et al. 2010; Mostek et al. 2013; Coupon et al. 2015; Durkalec et al. 2015b). Redshift-space galaxy clustering has also been measured to precisely constrain cosmological models, as exemplified by SDSS-III Baryon Oscillation Spectroscopic Survey (e.g., Guo et al. 2014; Alam et al. 2016; Okumura et al. 2016).

The halo occupation distribution (HOD) framework, which describes the average number of galaxies hosted by a halo as a function of halo mass, has been routinely used to describe the observed abundance and clustering of galaxies (e.g., Peacock & Smith 2000; Seljak 2000; Scoccimarro et al. 2001; Berlind & Weinberg 2002; Berlind et al. 2003; Kravtsov et al. 2004; Zheng et al. 2005). These observables constrain key quantities of the HOD, such as the minimum mass of halos that host at least one galaxy for a specific population. The HOD models have been applied to interpret galaxy clustering in the local Universe (e.g., Zehavi et al. 2004; Zheng et al. 2005, 2007; Cacciato et al. 2009, 2013; Leauthaud et al. 2011; Zehavi et al. 2011; More et al. 2015) and at higher redshifts up to $z \sim 7$ (e.g. , Brown et al. 2008; Wake et al. 2011; Geach et al. 2012; Durkalec et al. 2015a, 2015b; Harikane et al. 2016). However, most analyses of the HOD of high-redshift ($z \gtrsim 1$) galaxies have been conducted by using photometric samples (e.g., Wake et al. 2011; Coupon et al. 2012; Martinez-Manso et al. 2015; McCracken et al. 2015). Although imaging surveys provide large and deep samples at a much lower cost compared to spectroscopy, it is hard to resolve the redshift evolution, due to the contamination of back/foreground objects and associated large uncertainties in the redshift determination. For example, even for the estimates in the COSMOS catalogs based on 30-band photometry, the typical error is about $\Delta z/(1 + z) \sim 0.03$ (Ilbert et al. 2013; Laigle et al. 2016).

On the other hand, spectroscopic surveys at higher redshifts generally tend to be restricted to relatively bright, rare objects. Such samples may not represent the general galaxy population. Therefore, spectroscopic studies based on galaxies representative of the epoch's average population are required. It has been established that the bulk of star-forming galaxies follow a tight correlation between stellar mass ($M_*$) and star formation rate (SFR) over a wide range of redshift up to $z \sim 4$ (e.g., Daddi et al. 2007, 2009; Noeske et al. 2007; Whitaker et al. 2012; Kashino et al. 2013; Pannella et al. 2015) or even higher (e.g., Salmon et al. 2015); the so-called "main sequence". Galaxies along this sequence dominate the cosmic stellar mass density at $z \gtrsim 1$ (Ilbert et al. 2013), as well as the cosmic star formation rate density over cosmic history (Rodighiero et al. 2011; Schreiber et al. 2015). Therefore, main-sequence galaxies are representative of the average star-forming galaxy population at all redshifts.

In this work, we investigate the properties of halos that host galaxies representative of the main-sequence star-forming population, using spectroscopic selection of 516 objects at $1.43 \leqslant z \leqslant 1.74$ down to a stellar mass $10^{9.57} M_\odot$. This redshift range marks the transition epoch of cosmic star formation, from its peak to the presently continuing decline phase (Madau & Dickinson 2014). Galaxies in our sample are drawn from a near-IR spectroscopic campaign via the Fiber Multi-Object Spectrograph (FMOS) on the Subaru telescope, called the FMOS-COSMOS survey (Silverman et al. 2015). For this sample, we successfully map the small-scale structures below $1 h^{-1}$ comoving Mpc (cMpc), where the contribution from galaxies that reside in the same halo (i.e., one-halo term) becomes important. Such a high sampling rate per unit area is one of the unique advantages of our survey, which makes it complementary to the FastSound survey—another wide-field spectroscopic survey carried out with FMOS (Tonegawa et al. 2015; Okumura et al. 2016).

The paper is organized as follows. In Section 2, we provide an overview of the FMOS-COSMOS survey and describe our galaxy sample. We describe the methods employed to measure clustering in Section 3, and corrections for critical biases in Section 4. We present our clustering measurements in Section 5. Results are interpreted using an HOD modeling in Section 6, and we discuss the physical implications of the derived quantities in Section 7. We finally summarize our results and conclusions in Section 8. Throughout the paper, magnitudes are given in the AB system, and a flat $\Lambda$CDM cosmology with $(\Omega_m, \Omega_\Lambda) = (0.3, 0.7)$ is assumed. We express distances in comoving units and halo masses with the Hubble parameter $h$ in units of $100 \, \mathrm{km \, s^{-1} \, Mpc^{-1}}$, whereas stellar masses, and subsequently the stellar mass-to-halo mass ratios, are computed assuming $h = 0.7$. We use a Chabrier (2003) initial mass function (IMF). The conversion factor to a Chabrier IMF is $1/1.7$ for both stellar masses and SFRs given in our companion papers that use a Salpeter IMF (Kashino et al. 2013, 2017; Silverman et al. 2015). We use "log" to denote a logarithm with a base 10 ($\log_{10}$). For reference, Table 1 defines symbols used in this paper.

## 2. Data

### 2.1. Overview of the FMOS-COSMOS Survey

The galaxy sample used in this paper is constructed from the data set of the FMOS-COSMOS survey. Details of the observations, survey design, and data analysis are described elsewhere (Kashino et al. 2013; Silverman et al. 2015). FMOS is a near-infrared spectrograph on the Subaru telescope, with high multiplex capabilities; it allows the placement of 400 fibers, each with an aperture of $1\farcs2$ in diameter, over a circular field of 0.19 square degrees (Kimura et al. 2010). In cross-beam switching mode, roughly 200 objects can be observed at a

**Table 1**
Symbols Used in this Paper

| Symbol | Definition |
|---|---|
| $M_{\rm h}$ | halo mass |
| $M_*$ | stellar mass |
| $M_*^{\rm lim}$ | stellar mass limit for our sample ($10^{9.57}\,M_\odot$) |
| SFR | star formation rate |
| $f_{\rm H\alpha}^{\rm pre}$ | predicted H$\alpha$ flux (after dust obscuration) |
| $f_{\rm H\alpha}^{\rm lim}$ | lower limit of $f_{\rm H\alpha}^{\rm pre}$ for our sample ($1 \times 10^{-16}$ erg cm$^{-2}$ s$^{-1}$) |
| $f_{\rm fake}$ | fraction of misidentified objects in our sample |
| $\xi, \xi_{\rm dm}$ | real space correlation function of galaxies/dark matter |
| $w_{\rm p}, w_{\rm p,dm}$ | projected correlation function of galaxies/dark matter |
| $r_{\rm p}$ | comoving tangential separation |
| $\pi$ | comoving line-of-sight separation |
| $b$ | galaxy bias |
| $\omega(\theta)$ | angular correlation function |

given time. FMOS has an OH airglow suppression system that blocks strong OH emission lines (∼30% of the entire spectral window).

The survey is designed to detect H$\alpha$ emission lines with the *H*-long grating (1.6–1.8 $\mu$m, spectral resolution $R \approx 3000$). The H$\alpha$ and [N II]$\lambda\lambda$6548, 6584 lines are well-separated, and the accuracy of redshift determination is $\Delta z/(1 + z) = 2.2$e–4, corresponding to 66 km s$^{-1}$ (Silverman et al. 2015). We carried out additional observations with the *J*-long grating (1.11–1.35 $\mu$m) to detect the H$\beta$ and [O III]$\lambda\lambda$5007, 4959 emission lines, so as to establish the ionization conditions of $z \sim 1.6$ star-forming galaxies (Zahid et al. 2014; Kartaltepe et al. 2015; Kashino et al. 2017) and confirm redshift measurements for about half of the sample. Our survey covers the entire COSMOS field, with multiple FMOS footprints having some overlap regions. All data are reduced with the FMOS Image-Based Reduction Package (FIBRE-pac; Iwamuro et al. 2012).

### 2.2. Sample Selection

In this study, we use a catalog of galaxies with spectroscopic redshifts based on observations carried out from 2012 March to 2014 February, including 11 and 6 pointings in the central region of the COSMOS field with the *H*- and *J*-long gratings, respectively. The survey field, with an area of 0.810 deg$^2$, is shown in Figure 1. We remove the sky regions impacted by bright stars (shaded areas), which account for 4.9% of the survey area.

Our sample is selected from the COSMOS photometric catalog, which includes the Ultra-VISTA/VIRCAM photometry (McCracken et al. 2012; Ilbert et al. 2013). Photometric redshifts and stellar masses are derived by spectral energy distribution (SED) fitting with the Le Phare photometric code (Arnouts & Ilbert 2011), assuming Bruzual & Charlot (2003) stellar population synthesis models and a Chabrier IMF. The photometric catalog reaches a magnitude limit of $K_{\rm S} \approx 24$ mag, making it complete for more than 90% of the galaxies—down to a stellar mass $10^{9.57}\,M_\odot$ (Ilbert et al. 2013). We also derive star formation rates (SFRs) from SED fitting, assuming a constant star formation history—as extensively described in Silverman et al. (2015). The predicted H$\alpha$ flux ($f_{\rm H\alpha}^{\rm pre}$) for each galaxy is computed by a relation from Kennicutt (1998),
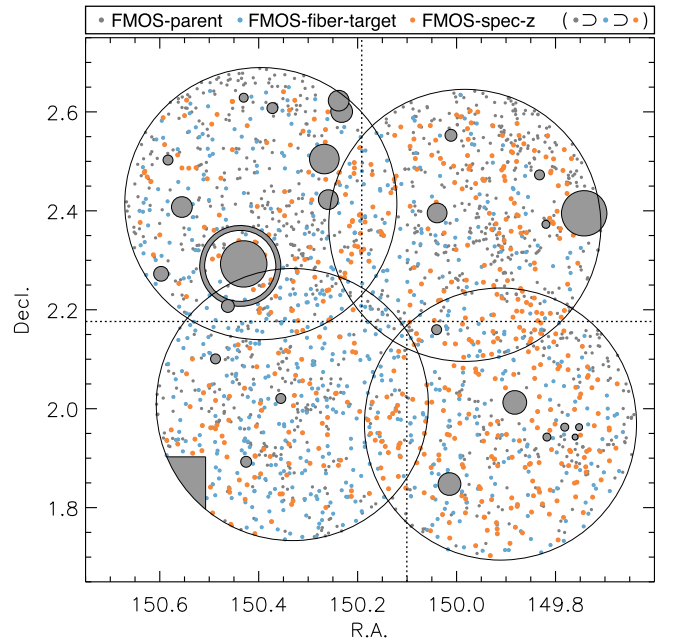


**Figure 1.** Four FMOS footprints (large circles) and galaxies in our FMOS samples (gray dots show the FMOS-parent sample; blue dots, the FMOS-fiber-target sample; orange dots, the FMOS-spec-$z$ sample). Dotted lines indicate partitions for jackknife resampling (see Section 3.3).

converted for use with a Chabrier IMF:

$$f_{\rm H\alpha}^{\rm pre} = \frac{1}{4\pi d_{\rm L}^2}\frac{{\rm SFR}(M_\odot\,{\rm yr}^{-1})}{4.65 \times 10^{-42}}10^{-0.4A_{\rm H\alpha}}, \qquad (1)$$

where $d_{\rm L}$ is luminosity distance. The values of $f_{\rm H\alpha}^{\rm pre}$ represent the total light coming from galaxies, as opposed to only that seen by a single fiber. Dust extinction of the stellar component $E_{\rm star}(B - V)$ is also derived from our SED fitting. This value has been converted to a nebular extinction to the H$\alpha$ emission as $A_{\rm H\alpha} = 3.325E_{\rm star}(B - V)/0.66$ (see, e.g., Kashino et al. 2013), assuming a Calzetti et al. (2000) extinction curve. For the remainder of the paper, these SED-based quantities are used for selection, whereas the spectroscopic redshifts and observed H$\alpha$ fluxes are incorporated into the final determination of their SFRs (Section 7.3). We describe in detail the steps of our sample selection below. Each subsample will be labeled at various stages of the selection, to make the treatment of biases clear.

Over an effective survey area (0.77 deg$^2$), we find 7006 galaxies in the COSMOS catalog that have stellar mass above a threshold mass $M_*^{\rm lim} \equiv 10^{9.57}\,M_\odot$ and photometric redshift ($z_{\rm phot}$) between 1.46 and 1.72. We refer to these mass (and photo-$z$) selected galaxies as $M_*$-*selected* sample. Our targets are restricted to those with the Ultra-VISTA/VIRCAM photometry $K_{\rm S} \leqslant 23.5$. We designate the subset of 6453 galaxies with both $M_* \geqslant M_*^{\rm lim}$ and $K_{\rm S} \leqslant 23.5$ as the $M_* + K_{\rm S}$-*selected* sample. In addition, we impose a limit on the predicted H$\alpha$ flux computed from the SED-based SFR and dust extinction (Equation (1)), as $f_{\rm H\alpha}^{\rm pre} \geqslant f_{\rm H\alpha}^{\rm lim} \equiv 1 \times 10^{-16}$ erg cm$^{-2}$ s$^{-1}$, to achieve an acceptable success rate of detecting the H$\alpha$ emission line with FMOS. This flux limit is equivalent to SFR $\approx 20\,M_\odot$ yr$^{-1}$ for galaxies with typical extinction $A_{\rm H\alpha} = 1$ mag at $z = 1.6$. Within the $M_* + K_{\rm S}$-selected sample, we find 2139 galaxies that have $f_{\rm H\alpha}^{\rm pre} \geqslant f_{\rm H\alpha}^{\rm lim}$, which are referred to as the *FMOS-parent* sample.

**Table 2**
Galaxy Samples

| Sample name | $N^a$ | $n$ ($10^{-3}\,h^3$ cMpc$^{-3}$) | Selection criteria | Mean $M_*{}^b$ | Median $M_*{}^b$ |
|---|---|---|---|---|---|
| $M_*$-selected | 7006 | 9.29$^c$ | $1.46 \leqslant z_{phot} \leqslant 1.72$, $M_* \geqslant 10^{9.57}\,M_\odot$ | 10.42 | 10.04 |
| $M_* + K_S$-selected | 6453 | 8.55$^c$ | $+ K_S \leqslant 23.5$ | 10.45 | 10.10 |
| FMOS-parent | 2319 | 3.07$^c$ | $+$ Predicted $f(H\alpha) \geqslant 10^{-16}$ erg s$^{-1}$ cm$^{-2}$ | 10.58 | 10.20 |
| FMOS-fiber-target | 1182 | ... | $+$ Observed | 10.59 | 10.23 |
| FMOS-spec-$z$ | 516 | 0.575$^d$ | $+$ H$\alpha$-detected at $1.43 \leqslant z \leqslant 1.74$ | 10.54 | 10.22 |

**Notes.**
[a] Number of galaxies in each sample.
[b] Mean and median stellar masses of each sample, expressed in $\log M_*/M_\odot$.
[c] Number density of photometrically selected sample, computed in the volume over $1.46 \leqslant z \leqslant 1.72$ ($7.54 \times 10^5\,h^{-3}$ cMpc$^3$).
[d] Number density of spectroscopic sample, computed in the volume over $1.43 \leqslant z \leqslant 1.74$ ($8.98 \times 10^5\,h^{-3}$ cMpc$^3$).

This sample is used as the input catalog for the process of allocating fibers to galaxies.

From the FMOS-parent sample, we observed 1182 galaxies with the *H*-long grating, which are referred to as *FMOS-fiber-target* sample. Of these, we select 516 galaxies that have a positive H$\alpha$ detection (S/N $\geqslant 1.5$) in the *H*-long window ($1.43 \leqslant z \leqslant 1.74$; median $z = 1.588$), to measure their spatial clustering. This final sample, referred to as the *FMOS-spec-z* sample, is restricted to galaxies having one or more emission lines with S/N $\geqslant 3$. In most cases (503/516), H$\alpha$ is detected at S/N $\geqslant 3$. For H$\alpha$ detections of low significance ($1.5 \leqslant$ S/N $< 3$), we require the detection of at least one other line (e.g., [O III]$\lambda5007$) at S/N $\geqslant 3$. Table 2 summarizes the size, number density, selection criteria, and mean/median stellar masses of each subsample. The total survey volume over the range of spectroscopic redshift ($1.43 \leqslant z \leqslant 1.74$) is $8.98 \times 10^5\,(h^{-1}\,\text{cMpc})^3$.

### 2.3. Sample Characteristics

In Figure 2, we show the SFR derived from SED, as a function of $M_*$, for galaxies in our sample. We note that stellar masses and SFRs of the FMOS-spec-z sample shown here are not revised by incorporating their spectroscopic redshifts, but rather the original values based on photometric redshifts used for the sample selection. In agreement with many studies, there is a clear correlation between SFR and $M_*$, i.e., the star-forming main sequence. It is evident that our sample traces the underlying distribution of star-forming galaxies. We find no significant difference in SFR between the FMOS-parent and FMOS-spec-z samples. However, it is shown that galaxies with relatively low SFRs (64%) are missed in the selection, due to our self-imposed limitation on the predicted H$\alpha$ flux. The median SFRs and 68th percentiles are shown in eight bins of stellar mass for both the $M_* + K_S$-selected (triangles) and FMOS-spec-z (blue squares) samples. The FMOS-spec-z sample is, on average, biased toward higher SFRs by 0.1–0.2 dex at $M_* \lesssim 10^{10.8}\,M_\odot$. However, such a bias is only about half the scatter in SFR of the $M_* + K_S$-selected sample. Therefore, we contend that the FMOS-spec-z sample is representative of the normal star-forming population at these redshifts.

We assess the completeness of our sample at a given stellar mass in Figure 3, which compares the stellar mass distributions of the $M_*$-selected, $M_* + K_S$-selected, and FMOS-parent samples. In the upper panel, the fractions of the number of galaxies in the $M_* + K_S$-selected or FMOS-parent samples to the number of galaxies in the $M_*$-selected sample are shown as
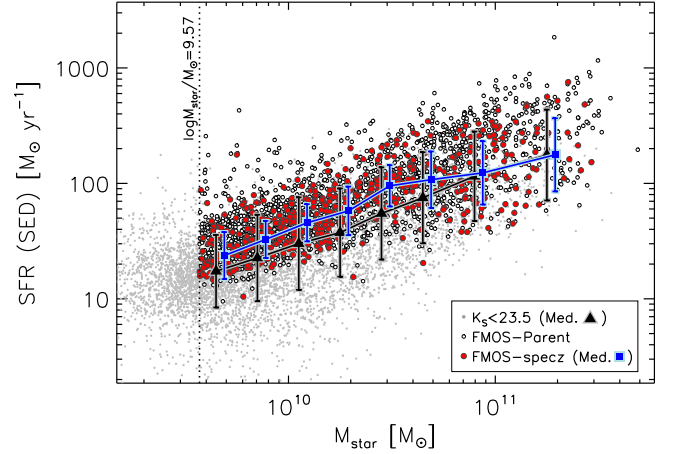


**Figure 2.** SFR (based on SED) vs. $M_*$ for galaxies in our sample. Gray dots show all galaxies with $K_S \leqslant 23.5$ and $1.46 \leqslant z_{phot} \leqslant 1.72$ in the FMOS survey area. The FMOS-parent and FMOS-spec-z samples are shown as open and red filled circles, respectively. The median SFRs in eight $M_*$ bins are indicated for the $M_* + K_S$-selected sample (filled triangles) and the FMOS-spec-z sample (filled blue squares). The error bars indicate the central 68th percentiles in SFR for each bin. A vertical dotted line indicates the threshold stellar mass ($M_*^{\rm lim} = 10^{9.57}\,M_\odot$) for our sample selection.
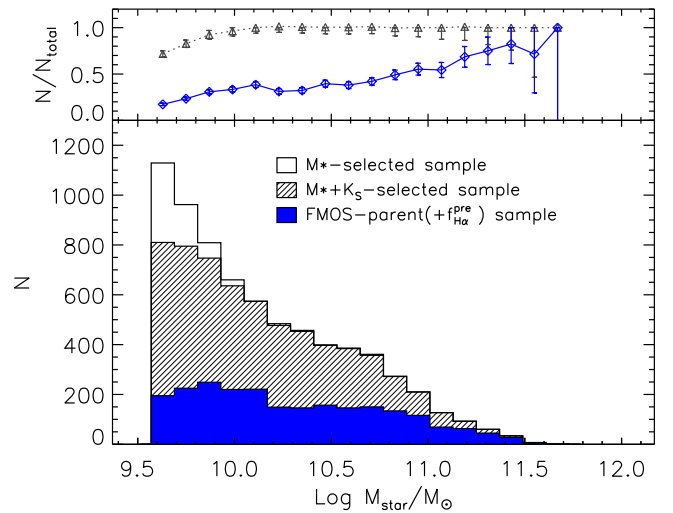


**Figure 3.** Stellar mass distributions of our pre-observation samples. Histograms show the numbers of galaxies in each 0.12 dex bin (empty histogram indicates $M_*$-selected; hatched histogram shows $M_* + K_S$-selected; filled histogram is the FMOS-parent sample). The upper panel shows the binned ratios of the number of galaxies in the $M_* + K_S$-selected sample (triangle) and FMOS-parent sample (diamond) to the number of galaxies in the $M_*$-selected sample, with error bars indicating the Poisson noise.
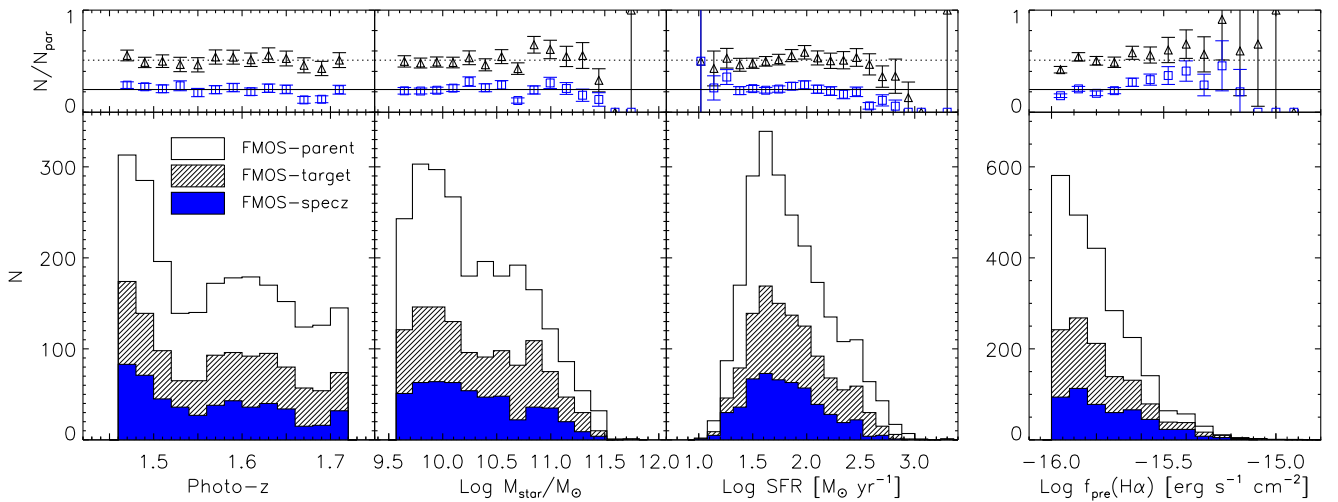
**Figure 4.** Distributions of photometric redshift, stellar mass, SFR (based on SED), and predicted Hα flux, from left to right, respectively: empty histograms indicate the FMOS-parent; hatched histograms show the FMOS-fiber-target; filled blue histograms are the FMOS-spec-$z$ sample. Upper panels show the binned ratios of the number of galaxies in the FMOS-fiber-target sample (triangles) or FMOS-spec-$z$ sample (squares) to the number of galaxies in the FMOS-parent sample, with error bars indicating the Poisson noise. Note: the left panel shows the distribution of photometric redshifts, even for the spec-$z$ sample.

a function of stellar mass. As evident, approximately 8% of galaxies in the $M_*$-selected sample are missed near the lower mass limit ($M_* \lesssim 10^{10}\,M_\odot$) by the inclusion of the $K_S \leqslant 23.5$ limit, while the stellar mass completeness reaches almost unity at $M_* > 10^{10.1}\,M_\odot$. In the FMOS-parent sample, a large fraction of galaxies in the $M_*$-selected sample are missed due to the limit on the predicted Hα flux. The sampling rate increases slowly with increasing stellar mass. As a consequence, the FMOS-parent sample has slightly higher mean and median stellar masses, as compared to the $M_*$-selected or $M_* + K_S$-selected sample (Table 2). Given an expected dependence of the clustering amplitude on stellar mass, such a selection bias may affect the observed correlation function. The effects of stellar mass incompleteness of our sample are further discussed, using mock samples, in Appendix B. We ensure that such effects do not significantly impact our results and conclusions.

We further evaluate the observational selection effects. Figure 4 shows the distribution of photometric redshift, stellar mass, SED-based SFR, and the predicted Hα flux for the FMOS-parent, -fiber-target, and -spec-$z$ samples. These values are not revised for spectroscopic redshifts. The fractions of the number of galaxies in the FMOS-fiber-target sample or in the FMOS-spec-$z$ sample to the number of galaxies in the FMOS-parent sample are presented in the upper panels. We highlight that the sampling rate is almost uniform as a function of photometric redshift, stellar mass, and SFR. As a consequence, the median values of these quantities of the FMOS-spec-$z$ sample are in good agreement with those of the FMOS-parent sample (see Table 2 for average masses). In contrast, the right panel of Figure 4 indicates that the FMOS-spec-$z$ sample is slightly biased toward having higher $f_{\mathrm{H}\alpha}^{\mathrm{pre}}$, as compared to the FMOS-parent sample. Such a bias is expected from the fact that stronger Hα lines are easier to detect, even though the median value of the predicted Hα flux of the FMOS-spec-$z$ sample ($\log f_{\mathrm{H}\alpha}^{\mathrm{pre}}/(\mathrm{erg\,s^{-1}\,cm^{-2}}) = -15.79$) is close to that of the FMOS-parent sample ($\log f_{\mathrm{H}\alpha}^{\mathrm{pre}}/(\mathrm{erg\,s^{-1}\,cm^{-2}}) = -15.83$). Therefore, we conclude that the observational sampling biases do not significantly affect the characteristics of the spec-$z$ sample, relative to the parent sample.

In Figure 5, we show the distribution of spectroscopic redshifts for 516 galaxies in the FMOS-spec-$z$ sample, with positions of OH lines highlighted. Wavelengths of OH lines are converted into redshifts based on the wavelength of the Hα emission line as $z = \lambda_{\mathrm{OH}}/\lambda_{\mathrm{H}\alpha} - 1$. It is evident that the number of successful detections is suppressed near OH contamination. We take into account these effects on the radial distribution of the target galaxies, as described in Section 3.2.

### 2.4. Mock Catalog

We validate our correction schemes (described in Section 4) for observational biases by using a set of mock samples constructed from cosmological $N$-body simulations. We utilize the new numerical galaxy catalog ($\nu^2$GC, Ishiyama et al. 2015) to take advantage of its large simulation volume. We use the medium-volume simulation ($\nu^2$GC-M) that was conducted by using $4096^3$ particles with a mass resolution of $2.2 \times 10^8\,h^{-1}\,M_\odot$ in a comoving box with a side length of $560\,h^{-1}$ cMpc. We employ the halo catalog at a scale factor $a = 0.384871$ ($z = 1.598$), close to our median redshift ($z = 1.588$), in which halos and subhalos are identified with the Rockstar algorithm (Behroozi et al. 2013b). For our purpose, the simulation box is divided into 64 subvolumes of $70 \times 70 \times 560\,(h^{-3}\,\mathrm{cMpc})^3$, each of which can enclose the entire FMOS survey volume. In each sub-box, we mimic our observations to construct realistic mock target catalogs. By doing so, the spatial distribution of mock galaxies reflects the artificial biases due to the fiber allocation and inhomogeneous detection, similar to the real data. We use them to establish and examine the correction schemes for these biases. Details of the construction of the mock catalogs and the correction schemes are described in Appendix A.

Furthermore, we assess the effect of stellar mass incompleteness. We use the Bolshoi simulation (Klypin et al. 2011) for this purpose, to take the advantage of its mass resolution (better than $\nu^2$GC) and the availability of merger histories of halos. This simulation traces $2048^3$ particles in a cubic box with a side length of $250\,h^{-1}$ cMpc. We use the public catalog at $a = 0.38435$ ($z = 1.602$) to construct mock samples. We
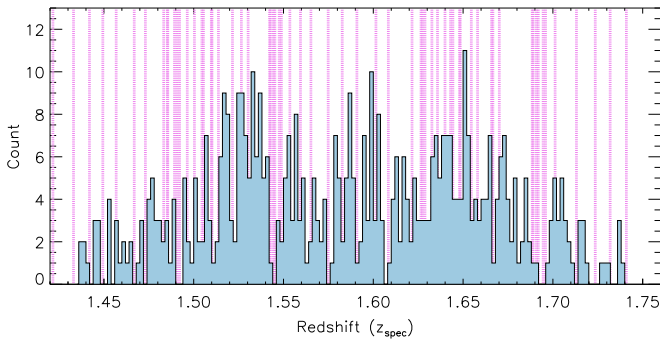
**Figure 5.** Distribution of the 516 spectroscopic redshifts from the FMOS-spec-z sample. Magenta stripes indicate positions of the OH airglow lines, which are shifted to the redshift of the H$\alpha$ emission line.

describe further the construction of the mock samples and the assessment of selection effects in Appendix B.

## 3. Clustering Measurement

### 3.1. Two-point Correlation Function

The two-point auto-correlation function is a powerful and commonly used tool to quantify the spatial distribution of galaxies. The real-space correlation function $\xi(r)$ measures the excess of the probability of finding pairs of galaxies as a function of their separation $r$. As a measure of the correlation function, we use the Landy & Szalay (1993) estimator:

$$\xi(r) = \frac{N_R(N_R - 1)}{N_D(N_D - 1)} \frac{DD(r)}{RR(r)} - \frac{N_R - 1}{N_D} \frac{DR(r)}{RR(r)} + 1, \quad (2)$$

where $N_D$ and $N_R$ are the numbers of galaxies and random objects, respectively; $DD(r)$, $DR(r)$, and $RR(r)$ are the numbers of data–data, data–random, and random–random pairs with a comoving separation within the interval $[r, r + dr]$, respectively.

The radial distance computed from the redshift is different from the actual distance, due to the peculiar motion of the galaxy, resulting in distortion of the correlation function measured in redshift space. To minimize these effects, we use the standard practice for clustering analysis of measuring the correlation function on a two-dimensional grid, parallel and perpendicular to the line of sight, and integrating along the line-of-sight direction. We define the separation of objects following Fisher et al. (1994). Given a pair of objects at positions $r_1$ and $r_2$ in the redshift comoving space, the separation $s$ and line-of-sight vector $l$ are defined as $s = r_1 - r_2$ and $l = (r_1 + r_2)/2$, respectively. The parallel ($\pi$) and perpendicular ($r_p$) separations are given, respectively, as

$$\pi \equiv |s \cdot l| / |l|, \qquad r_p \equiv \sqrt{|s|^2 - \pi^2}. \quad (3)$$

The *projected* correlation function $w_p(r_p)$ is related to the real-space correlation function as follows:

$$w_p(r_p) = 2 \int_0^{\pi_{\max}} \xi(r_p, \pi) d\pi. \quad (4)$$

If $\pi_{\max}$ is infinity, the redshift space distortions have no effect on $w_p(r_p)$. In practice, it should be large enough to eliminate the effect of peculiar motions, but finite to avoid adding noise to the measurements. We adopt $\pi_{\max} = 30\,h^{-1}$ cMpc, which corresponds to $\Delta z \approx \pm 0.024$ or $\Delta v \approx \pm 2800\,\mathrm{km\,s^{-1}}$. We

count pairs of galaxies and random objects on the two-dimensional grid, which is binned logarithmically in the $r_p$ direction and linearly in $\pi$. Considering the small number of galaxy pairs with a small separation ($r_p \lesssim 1\,h^{-1}$ cMpc), we employ a set of variable-size bins that have larger widths at small scales.

### 3.2. Construction of Random Samples

To measure galaxy clustering with the estimator given in Equation (2), we need to construct a reference random sample that follows the same geometrical properties as the real data. To avoid introducing shot noise, the random sample contains a large number of objects ($N_R = 45,000$ for our case), which is about 90 times larger than the spec-z sample. Random objects are distributed uniformly across the effective survey area, but we consider a non-uniform radial distribution to reflect the realistic redshift distribution of our sample.

Determination of the radial distribution of observed galaxies to create the random sample represents a significant challenge. The simplest way is to randomly assign objects the same redshifts as the sample galaxies. However, this method is useful only for wide-field surveys that are not impacted strongly by cosmic variance. Given the survey area, our spectroscopic redshifts may suffer from this effect, as the measured redshift distribution may reflect specific structures in the galaxy distribution. It can lead to an artificial line-of-sight clustering in the random catalog and subsequent underestimation of the correlation strength of real galaxies. In fact, with a reference catalog constructed in this way, we find a smaller correlation length (by $\sim 0.6\,h^{-1}$ cMpc) than our final result obtained with the random catalog constructed as described below.

To define the underlying redshift distribution of galaxies, we use $\sim 10^4$ galaxies with $K_S \leqslant 23.5$ and $M_* \geqslant M_*^{\mathrm{lim}}$ in the COSMOS photometric catalog (Ilbert et al. 2013). Here, no limitation on the photometric redshift is applied. Figure 6(a) shows the distribution of the radial comoving distance calculated from the photometric redshift of those galaxies. We smooth the binned distribution (histogram) with a Gaussian kernel with a standard deviation of 150, 250, or 450 $h^{-1}$ cMpc. The distribution smoothed with the shortest kernel (150 $h^{-1}$ cMpc) still traces specific structures (red line in Figure 6(a)). In contrast, smoothing with the longest kernel (450 $h^{-1}$ cMpc) may produce a slight artificial enhancement at both ends and suppression around the peak of the distribution (blue dashed line). Therefore, we decided to use the distribution smoothed with an intermediate scale of 250 $h^{-1}$ cMpc (green dotted line), which traces well the global shape of the distribution. We then estimate the intrinsic distribution of true redshifts of the FMOS-parent sample by taking into account the uncertainties on the photometric redshifts. Figure 6(b) illustrates our method. We extract the smoothed distribution between $1.46 \leqslant z_{\mathrm{phot}} \leqslant 1.72$ (dotted line), and convolve it with a Gaussian function that has a standard deviation $\sigma_{z_{\mathrm{phot}}} = 0.062$, which is a typical error on the photometric redshift of our sample galaxies. As a result, we obtain the realistic radial distribution of our parent sample (dashed line).

We also consider the effects of the OH airglow mask and the inhomogeneous sensitivity. The noise level of pixels impacted by the sky contamination is much higher than the typical level of $\sim 5 \times 10^{-19}$ erg cm$^{-2}$ s$^{-1}$ Å$^{-1}$ (see Figure 11 in Silverman et al. 2015); thus, these pixels are ignored in the emission-line fitting. As a consequence, the detection rate decreases at the
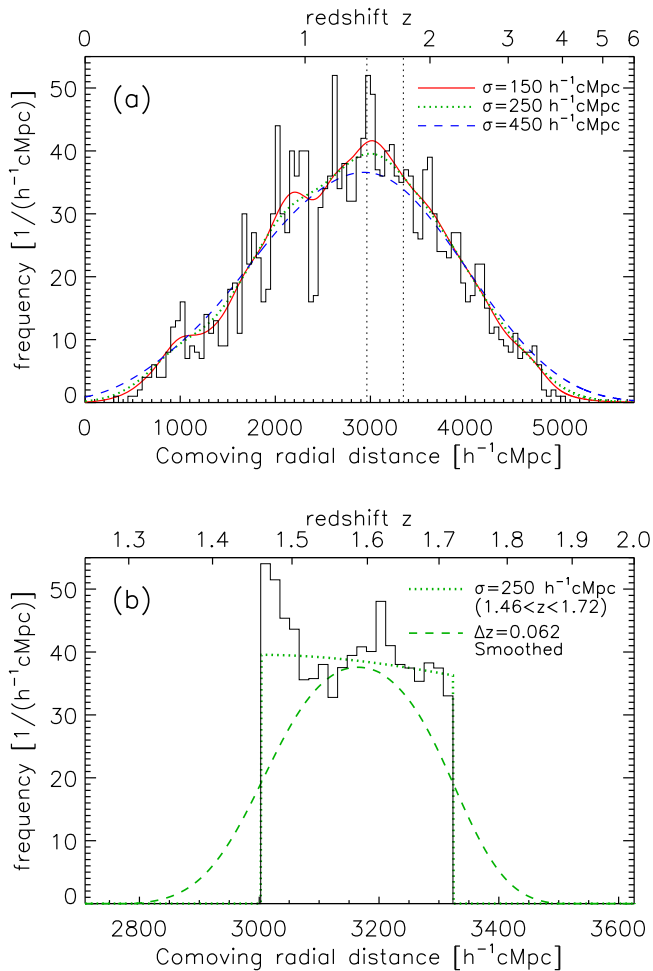
**Figure 6.** Distribution of comoving radial distance (and corresponding redshift) for galaxies with $K_S \leqslant 23.5$ and $M_* \geqslant M_*^{\mathrm{lim}}$ within the COSMOS photometric catalog (Ilbert et al. 2013). Panel (a): the solid-line histogram indicates the distribution of all galaxies. Solid, dotted, and dashed curves show the smoothed distribution with a Gaussian kernel of $\sigma = 150$, 250, and 450 $h^{-1}$ cMpc, respectively. Vertical dotted lines indicate the photometric redshift range for the FMOS–parent sample ($1.46 \leqslant z \leqslant 1.72$). Panel (b): a solid-line histogram and dotted curve are extracted from those in the top panel within the redshift range of $1.46 \leqslant z \leqslant 1.72$. The dashed curve shows the expected underlying distribution of the true redshift of our parent sample, which is determined by convolving the dotted curve with a Gaussian kernel with $\sigma_z = 0.062$.

positions around the OH lines (Figure 5). We assess the detection rate of the H$\alpha$ emission line as a function of redshift over the FMOS *H*-long spectral window by performing a set of Monte-Carlo simulations. For each of 516 galaxies in the FMOS-spec-*z* sample, we create artificial spectra with a multi-Gaussian profile that has the measured amplitude and line width for H$\alpha$ and [N II]$\lambda\lambda$6548, 6583. Gaussian noise is added to these artificial spectra based on the noise spectrum of each galaxy. Here, the noise level is intentionally increased by a factor of 1.25 to account for the fact that the noise level tends to be underestimated relative to the actual pixel variance (see Silverman et al. 2015). We then perform a fitting procedure for these artificial spectra in the same manner as the data and examine whether the H$\alpha$ line is recovered. We assess the detectability of this artificial signal at all pixels in the *H*-long window by scanning the entire spectral range. We define successful cases to be those for which the line is detected with S/N $\geqslant 3$ and the difference between the input and measured redshifts $\Delta z/(1 + z) < 0.001$. Figure 7 shows the average

detection rate, with positions of OH lines marked. As is evident, a decrease of the detection rate is seen at the positions of the OH lines. In addition, the detectability falls down at both ends of the spectral coverage, where the noise level becomes relatively high due to the instrumental characteristics. These features are expected to introduce artificial clustering in the line-of-site direction. This can be cancelled by using the random catalog that includes the same features. We take into account these instrumental effects on the radial selection function in the random catalog by multiplying the intrinsic redshift distribution (derived above, based on the photometric catalog—indicated by the dashed curve in Figure 6(b)) by this weight function shown in Figure 7. The constructed *final* weight function is given in Figure 8, in comparison with the distribution of observed spectroscopic redshifts. The global trend of the data is well-represented by the weight function, although some spikes in the histogram may reflect the intrinsic clustered structures of galaxies. We demonstrate the impact of this inhomogeneous selection function and the validity of our correction scheme in Appendix A.

### 3.3. Statistical Error Estimates

We estimate the statistical errors of the observed correlation function using a jackknife resampling method. We construct twenty jackknife samples of the FMOS-spec-*z* sample as follows. We first divide the entire survey area into four contiguous subregions of equal area (see Figure 1). The survey volume is then divided into five slices along the line of sight. Each subvolume has a typical transverse side length of $20\,h^{-1}$ cMpc and a depth of $80\,h^{-1}$ cMpc. For each jackknife sample, a subvolume (1/20 of the full volume) is omitted in turn.

Because each spatial structure of galaxies affects the pair counts at different separations, the values of $w_{\mathrm{p}}(r_{\mathrm{p}})$ at different $r_{\mathrm{p}}$'s are correlated. Therefore, we need to use the full covariance matrix to fit a model to the data. The associated covariance matrix $\boldsymbol{C}_{ij}$ is estimated from the jackknife samples as follows:

$$\boldsymbol{C}_{ij} = \frac{N-1}{N} \sum_{k=1}^{N} [X_i^k - \langle X_i \rangle][X_j^k - \langle X_j \rangle], \qquad (5)$$

where $N = 20$ is the number of jackknife samples, $X_i^k$ is the projected correlation function at the *i*th separation measured for the *k*-th jackknife sample, and $\langle X_i \rangle$ is the average of $X_i^k$ from $k = 1$ to $N$. Here, we use the logarithmic values $\log w_{\mathrm{p}}(r_{\mathrm{p}})$, rather than $w_{\mathrm{p}}(r_{\mathrm{p}})$, following the suggestion of Norberg et al. (2009). The choice of logarithmic or linear value does not change our conclusions.

Each element of the covariance matrix has large uncertainties, due to the small number of jackknife samples. Therefore, we separately smooth the covariance matrix for the diagonal or off-diagonal elements, following Mandelbaum et al. (2013). The diagonal elements (i.e., variance) are smoothed by a center-weighted kernel as follows:

$$(\sigma_i^2)^{\mathrm{smooth}} = (\boldsymbol{C}_{i-1,i-1} + 2\boldsymbol{C}_{ii} + \boldsymbol{C}_{i+1,i+1})/4. \qquad (6)$$

Figure 9 (top panel) shows the jackknife covariance diagonal elements and the smoothed values. The discontinuities at $r_{\mathrm{p}} \approx 0.15$ and $6\,h^{-1}$ cMpc are effectively mitigated, but the global shape is preserved. To smooth the off-diagonal
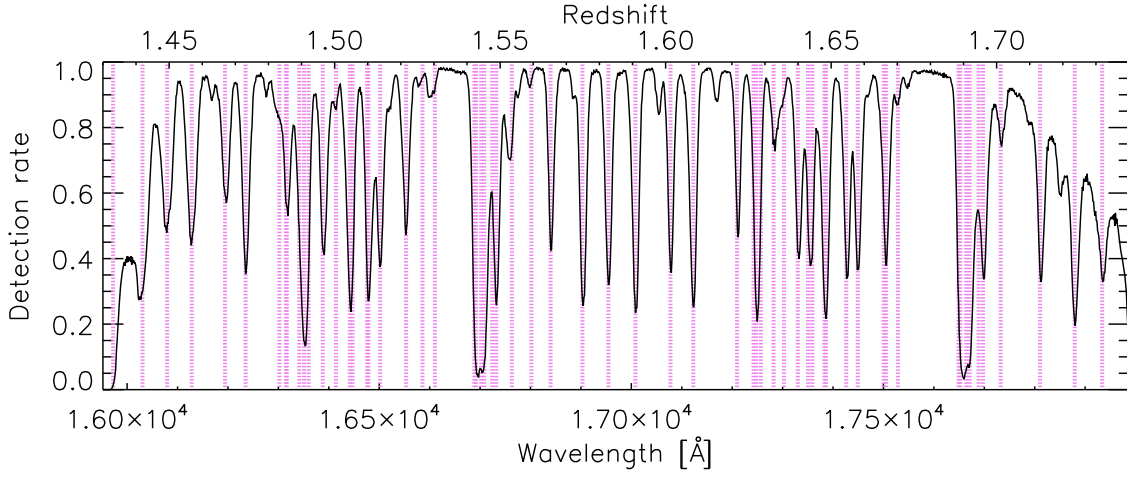
**Figure 7.** Average detection rate of the Hα emission line at each observed wavelength, with corresponding redshift on the upper axis. Magenta stripes indicate the positions of the OH airglow lines. The detection rate decreases in the vicinity of the OH lines and toward the ends of the spectral coverage, due to relatively higher noise level.
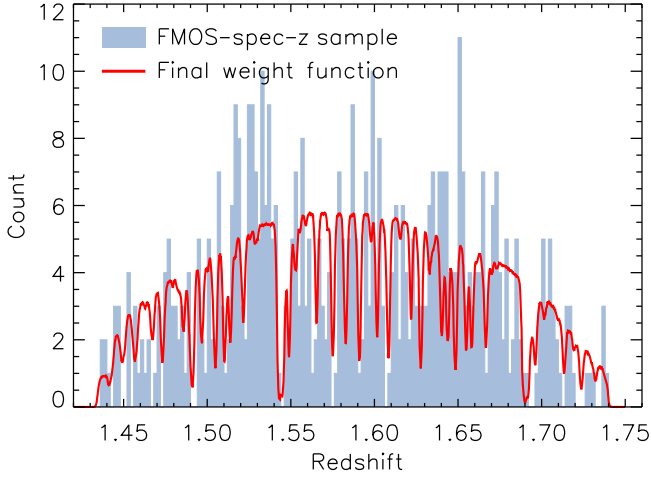


**Figure 8.** Comparison between the distribution of observed spectroscopic redshift (filled histogram) and the line-of-sight weight function for the random catalog (red line). This final weight function is the product of the smooth distribution (dashed line in Figure 6(b)) and weight for the inhomogeneous detectability of Hα (Figure 7).

elements, we first define the correlation matrix $R_{ij} = C_{ij}/\sqrt{C_{ii}C_{jj}}$, then smooth it to obtain $R_{ij}^{\mathrm{smooth}}$ with a $3 \times 3$ kernel, as follows:

$$\frac{1}{16}\begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix}. \quad (7)$$

The final smoothed covariance matrix $C_{ij}^{\mathrm{smooth}}$ is calculated as:

$$C_{ij}^{\mathrm{smooth}} = \begin{cases} (\sigma_i^2)^{\mathrm{smooth}} & (i = j), \\ R_{ij}^{\mathrm{smooth}}\sqrt{(\sigma_i^2)^{\mathrm{smooth}}(\sigma_j^2)^{\mathrm{smooth}}} & (i \neq j). \end{cases} \quad (8)$$

The original and smoothed correlation matrices are shown in Figure 9 (middle and bottom panels). The majority of pixel-to-pixel fluctuations are eliminated well by smoothing, but the overall trend is retained. Finally, we note that our conclusion does not depend on whether we are using the full covariance matrix or only the diagonal elements, or details of the smoothing method.

### 3.4. Model Fit

To obtain physical insight, we fit a power-law model, a biased dark matter model, and an HOD model to the observed correlation function. We use 11 data points at $-1.5 < \log(r_{\mathrm{p}}/h^{-1}\,\mathrm{cMpc}) < 1.1$ for the model fitting. At larger scales $(r_{\mathrm{p}} \gtrsim 15\,h^{-1}\,\mathrm{cMpc})$, the contribution from the integral constraint becomes greater than 10% of the measured clustering amplitude (see Section 4.2).

We define the posterior distribution of our model parameters to be given by

$$\mathcal{P} \propto \mathcal{P}_{\mathrm{prior}}(\boldsymbol{p})\exp\left(-\frac{\chi_{w_{\mathrm{p}}}^2 + \chi_{n_{\mathrm{tot}}}^2}{2}\right), \quad (9)$$

where $\mathcal{P}_{\mathrm{prior}}(\boldsymbol{p})$ is a prior probability distribution for a set of model parameters $\boldsymbol{p}$. In the brackets, $\chi_{w_{\mathrm{p}}}^2$ is calculated from the observed correlation function as

$$\chi_{w_{\mathrm{p}}}^2 = \sum_{i=0}^{N}\sum_{j=0}^{N}[X_i^{\mathrm{mod}} - X_i^{\mathrm{obs}}](C_{ij}^{\mathrm{smooth}})^{-1}[X_j^{\mathrm{mod}} - X_j^{\mathrm{obs}}], \quad (10)$$

where $X_i^{\mathrm{mod}}$ and $X_i^{\mathrm{obs}}$ are logarithms of the model and observed correlation functions at the $i$th separation, $N$ is the number of data points, and $(C_{ij}^{\mathrm{smooth}})^{-1}$ is the inverse of the smoothed covariance matrix defined in Section 3.3. To compare with the measurements based on pair counting in bins of $r_{\mathrm{p}}$, we calculate the average of the model $\log w_{\mathrm{p}}(r_{\mathrm{p}})$ in each $r_{\mathrm{p}}$ bin $[r_{\mathrm{min}} : r_{\mathrm{max}}]$ as follows:

$$X_i^{\mathrm{mod}} = \log\left[2\int_{r_{\mathrm{min}}}^{r_{\mathrm{max}}} w_{\mathrm{p}}^{\mathrm{mod}}(r_{\mathrm{p}})r_{\mathrm{p}}dr_{\mathrm{p}}/(r_{\mathrm{max}}^2 - r_{\mathrm{min}}^2)\right]. \quad (11)$$

In Equation (9), the latter term $\chi_{n_{\mathrm{tot}}}^2$ denotes the contribution from the constraint on galaxy abundance, which is considered only for HOD modeling (Section 6.1). To sample the posterior distribution of our parameters, we adopt a Markov Chain Monte Carlo (MCMC) technique, using the software *emcee* (Foreman-Mackey et al. 2013). We analyze a chain of 152400 steps that follows 25400 burn-in steps to find the parameter set that provides the maximum posterior probability, and to
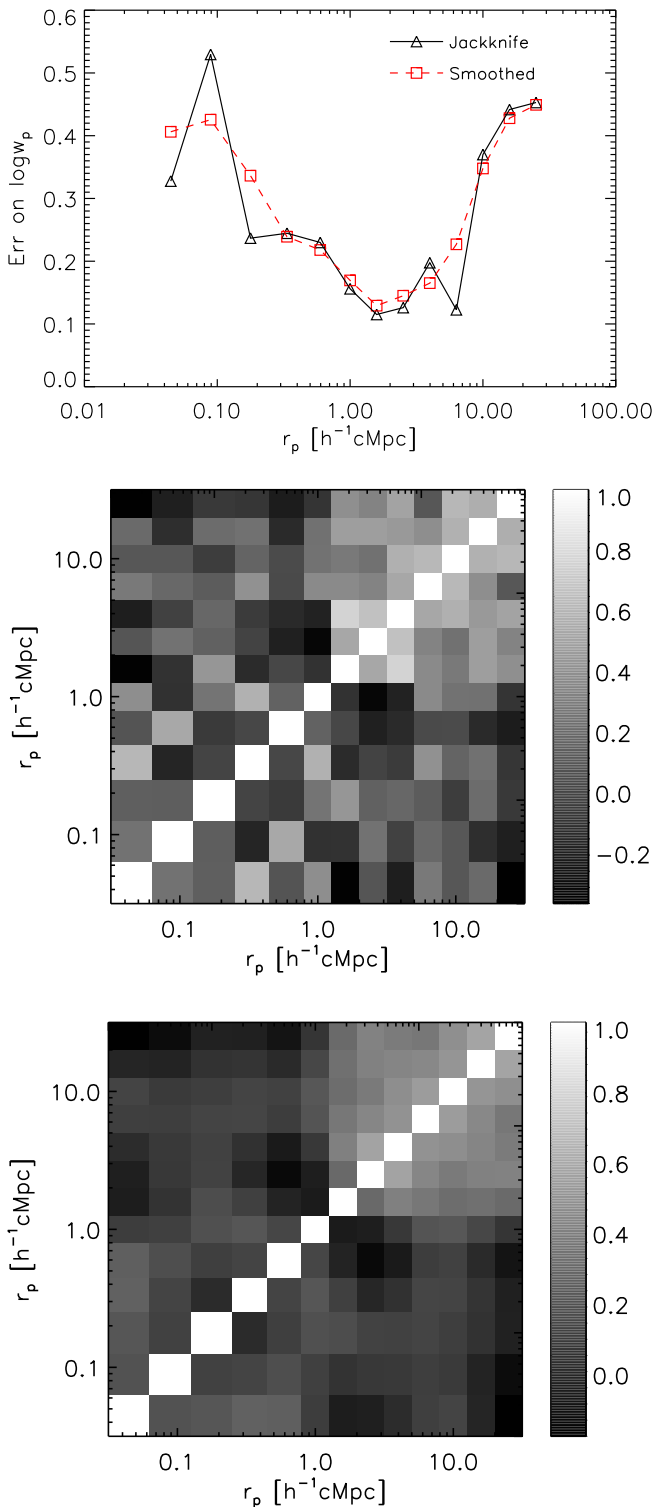
**Figure 9.** Top panel: errors on $\log w_{\rm p}(r_{\rm p})$ (square root of the diagonal term of the covariance matrix $C_{ij}$), derived from the jackknife resampling (triangles and solid line). The smoothed errors are indicated by squares and a dashed line. Middle panel: jackknife correlation matrix for $\log w_{\rm p}$. Bottom panel: smoothed correlation matrix.

evaluate the posterior probability distribution of each parameter.

## 4. Corrections for Critical Biases

In spectroscopic galaxy surveys, various observational effects can cause non-negligible, artificial biases. As described

in Section 3.2, we deal with the effects of the non-uniform detection of the H$\alpha$ emission line along the redshift direction by using a modified random sample (see also Appendix A). Here, we describe our treatment for other systematic effects.

### 4.1. Fiber Allocation

The most important issue is the impact of fiber allocation, which artificially distorts the on-sky distribution of objects, if not all galaxies in the input catalog are observed. For our FMOS observations, galaxies are selected from an input catalog by using the Echidna Spine-to-Object allocation software (Akiyama et al. 2008) to maximize operational efficiency. The FMOS fibers are uniformly embedded in the field-of-view, and they can move within a limited circular patrol area of 174 arcsec in diameter. Once a pair of fibers is allocated for one galaxy, the opportunity for its neighboring galaxies to be observed at the same time decreases due to the lack of fibers and/or the avoidance of fiber entanglement, although the patrol areas of adjacent fibers overlap with each other. As a consequence, the sampling rate of close galaxy pairs is suppressed at scales less or similar to the minimum separation of fibers ($\sim 1'.6$). In addition, the sampling rate is different across the survey area, due to different number of repeated exposures covering the same footprint and the overlapping regions. These characteristics of our observations affect the observed correlation function of galaxies, and thus need to be properly removed to measure the galaxy clustering.

We correct the observed correlation function for these biases by using a simple weighting scheme, in which each galaxy-galaxy pair is weighted in response to their angular separation (de la Torre et al. 2011; Durkalec et al. 2015b). The weight is defined as a ratio of the probability of finding pairs with a given angular separation in the input catalog to the sample of galaxies that were assigned to be observed. The weight function can be expressed by the angular correlation function of these two samples ($\omega_{\rm par}(\theta)$ and $\omega_{\rm tar}(\theta)$, respectively) as follows (e.g., Hawkins et al. 2003):

$$f(\theta) = \frac{1 + \omega_{\rm par}(\theta)}{1 + \omega_{\rm tar}(\theta)}. \tag{12}$$

It may be straightforward to determine the weight function based on the real data (i.e., the FMOS-parent and FMOS-fiber-target samples). However, the statistical uncertainties in the resulting weight function become large because the sample size is insufficient. Therefore, we decided to use mock samples to avoid such large statistical uncertainties. In Appendix A, we fully describe the construction of these mock samples and determination of the weight function, then demonstrate the validity of our correction scheme.

### 4.2. Integral Constraint

Due to the finite survey area, the observed correlation function is underestimated by a scale-independent constant value $C$, which is known as the integral constraint:

$$w_{\rm p}(r_{\rm p}) = w_{\rm p}^{\rm obs}(r_{\rm p}) + C. \tag{13}$$

First, we calculate the integral constraint for the real-space correlation function $\xi(r)$, following Roche et al. (1999), as

$$C_\xi = \frac{\sum_i \xi^{\mathrm{mod}}(r_i) RR(r_i)}{\sum_i RR(r_i)}, \qquad (14)$$

where $RR(r_i)$ is the number of random-random pairs whose separation is in the linearly spaced $i$th bin (bin size $\Delta r = 1\,h^{-1}$ cMpc). The summation is taken over the entire survey volume. The relation between $C$ and $C_\xi$ is simply given by

$$C = 2 \int_0^{\pi_{\mathrm{max}}} C_\xi\, d\pi = 2 C_\xi \times 30\,h^{-1}\,\mathrm{cMpc}. \qquad (15)$$

We use a biased nonlinear correlation function of dark matter as the model function, i.e., $\xi^{\mathrm{mod}} = b^2 \xi_{\mathrm{dm}}$. First we evaluate $b^2$ by comparing the nonlinear projected correlation function of dark matter with the observed $w_{\mathrm{p}}^{\mathrm{obs}}(r_{\mathrm{p}})$, then calculate the integral constraint using Equations (14) and (15). Next, we revise $b^2$, and recalculate $C$. By repeating this process until convergence, we find $C = 0.80\,h^{-1}\,$cMpc for our data. This value is comparable to the observed correlation amplitude at scales greater than $r_{\mathrm{p}} \sim 20\,h^{-1}$ cMpc. Therefore, we use only measurements at scales smaller than this for analyses.

Mock catalogs can be used to estimate the integral constraint. We construct a series of mock samples (Mock-par-spec$z$ samples; see Appendix A.1), where the sky coverage of the survey is replicated, and compare the average amplitude of their correlation functions, which are measured with an appropriate random catalog, to the measurement obtained from the entire simulation box which corresponds to a sky coverage of $10 \times 10\,\mathrm{deg}^2$. The suppression of the correlation function is found to be $C = \delta\, w_{\mathrm{p}} \approx 2\,h^{-1}\,$cMpc. This ensures that the estimate above is reasonable and that the effect is negligibly small at $r_{\mathrm{p}} \lesssim 10\,h^{-1}$ cMpc.

### 4.3. Contamination by Fake Detection

We correct the observed correlation function for contamination by fake sources for which a spurious signal or non-H$\alpha$ emission line is misidentified as H$\alpha$. Such contamination is expected to reduce the amplitude of the observed correlation function since these sources are not correlated with other real galaxies. Assuming that such sources are randomly distributed over the survey volume, the correlation function needs to be corrected by increasing its amplitude by a factor $1/(1 - f_{\mathrm{fake}})^2$ where $f_{\mathrm{fake}}$ is the fraction of fake sources in the sample.

We assess the reliability of our redshift measurements using an independent spectroscopic survey that includes the same galaxies. In the FMOS-spec-$z$ sample, we find 28 galaxies that have a robust redshift measurement (confidence class 3 or 4) from the zCOSMOS-deep survey (Lilly et al. 2007). Of these, 24 galaxies have redshift measurements from the FMOS-COSMOS and zCOSMOS surveys that are mutually consistent ($\Delta z < 0.01$), but four other objects have inconsistent measurements. Assuming that our FMOS measurements are wrong for these objects, but all 28 zCOSMOS redshifts are correct, we find a plausible fraction of fake detection to be $f_{\mathrm{fake}} = 0.14$. This fraction corresponds to a 26% underestimate of the correlation function. We note that the four objects with a possible wrong measurement are neither revised nor removed from the sample for analysis, although one of

them matches the zCOSMOS measurement if we suppose that the [O III]$\lambda$5007 is misidentified as H$\alpha$.

In the fitting process, we use the observed correlation function without the correction for the contamination from fake detections. Instead, we include this effect in our analyses by handling $f_{\mathrm{fake}}$ as a parameter to be estimated along with other model parameters. Namely, we compare the model multiplied by $(1 - f_{\mathrm{fake}})^2$ to the data. Assuming Poisson statistics, 4/28 approximately translates to $0.14 \pm 0.06$, which is imposed as a prior probability distribution on $f_{\mathrm{fake}}$. However, the zCOSMOS-deep survey is targeting rather higher redshift galaxies at $z > 2$, but has a lower sampling rate at the redshift range of our FMOS sample. Therefore, the fake fraction may be overestimated. Considering the large uncertainty in the estimation of $f_{\mathrm{fake}}$, we report results obtained by using this prior on $f_{\mathrm{fake}}$ as fiducial, but present results with $f_{\mathrm{fake}}$ fixed to zero for reference. In all cases, our results for these two cases are consistent at a $1\sigma$ level. We note that this prior dominates the posterior probability distribution of $f_{\mathrm{fake}}$ for the power-law and biased dark matter models (see Sections 5.1 and 5.2, respectively) because there is no other information that can constrain the amplitude of the correlation function. This prior on $f_{\mathrm{fake}}$ broadens the posterior distribution of the parameter that determines the correlation amplitude, i.e., correlation length or galaxy bias. In contrast, in the HOD modeling (Section 6.1), the posterior of $f_{\mathrm{fake}}$ is modified from the prior as the amplitude and shape of correlation function, and the total galaxy abundance are linked through the model.

## 5. Results

The projected correlation function $w_{\mathrm{p}}(r_{\mathrm{p}})$ is computed for a sample of 516 star-forming galaxies at $1.43 \leqslant z \leqslant 1.74$ in the central $0.81\,\mathrm{deg}^2$ (effectively $0.77\,\mathrm{deg}^2$) of the COSMOS field. Figure 10 presents the observed $w_{\mathrm{p}}(r_{\mathrm{p}})$ with and without correction for the scale-dependent effect of fiber allocation. In both cases, the measurements are corrected for the integral constraint, but not for the effect of fake detections. Instead, the model functions are reduced by $(1 - f_{\mathrm{fake}})^2$, where $f_{\mathrm{fake}} = 0.14$. As is evident, the amplitude of the correlation function without correction for the fiber allocation effect is slightly suppressed, as compared to the corrected values at small scales below $r_{\mathrm{p}} \sim 1\,h^{-1}$ cMpc (see Section 4.1 and Appendix A for details). We use the corrected values as fiducial measurements throughout this paper. The error bars indicate the standard deviation of $\log w_{\mathrm{p}}(r_{\mathrm{p}})$ estimated from jackknife resampling (see Section 3.3).

### 5.1. Power-law Model

It is known that a real space galaxy correlation function $\xi(r)$ can be described well by a power-law function as $\xi(r) = (r/r_0)^{-\gamma}$ (e.g., Totsuji & Kihara 1969; Zehavi et al. 2002), where $r_0$ and $\gamma$ are a correlation length and power-law slope, respectively. The correlation length denotes how strongly galaxies are clustered. With this form of $\xi(r)$, from Equation (4), $w_{\mathrm{p}}(r_{\mathrm{p}})$ can be expressed as

$$w_{\mathrm{p}}(r_{\mathrm{p}}) = r_{\mathrm{p}} \left( \frac{r_{\mathrm{p}}}{r_0} \right)^{-\gamma} \frac{\Gamma\left(\frac{1}{2}\right) \Gamma\left(\frac{\gamma-1}{2}\right)}{\Gamma\left(\frac{\gamma}{2}\right)}, \qquad (16)$$

where $\Gamma$ is Euler's Gamma function and the integration limit $\pi_{\mathrm{max}}$ of Equation (4) is taken as infinity. We fit the form of
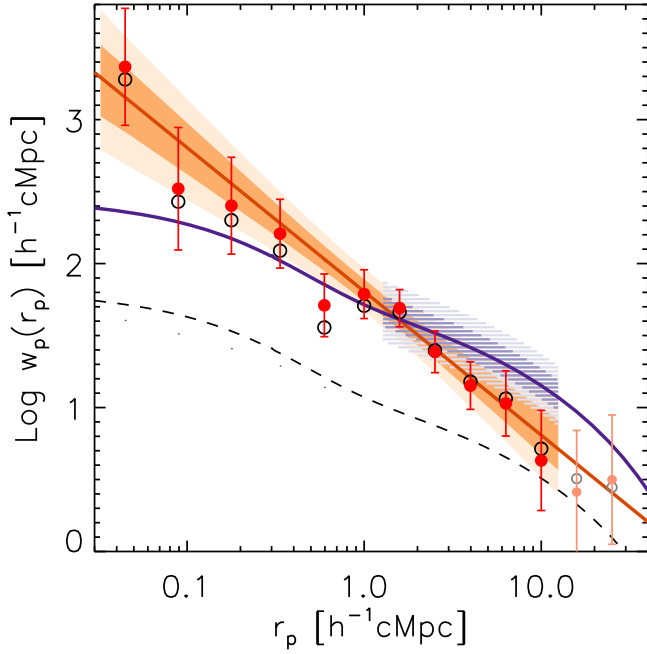
**Figure 10.** Projected two-point correlation function $w_p(r_p)$ of our galaxy sample at $1.43 \leqslant z \leqslant 1.74$. Filled red circles show the fiducial measurements of $w_p(r_p)$, corrected for the fiber allocation effects, whereas open circles indicate the uncorrected $w_p(r_p)$. The error bars indicate $1\sigma$ uncertainties (see Section 3.3). The best-fit power-law and biased dark matter models are shown by orange and purple lines, respectively. The 68% and 95% confidence intervals of each model are shown by dark and light shaded regions. Note that the observed $w_p(r_p)$ is not corrected for suppression of the amplitude due to fake detections; the best-fitting models are reduced by multiplying by a factor of $(1 - f_{fake})^2$, where $f_{fake} = 0.14$. The dashed line indicates the correlation function of dark matter ($b = 1$).

Equation (16), multiplied by the contamination factor $(1 - f_{fake})^2$, to the observed $w_p(r_p)$ (see Figure 10), then search the parameter space ($r_0$, $\gamma$, $f_{fake}$) with the MCMC procedure. We find a correlation length to be $r_0 = 5.26^{+0.75}_{-0.62} \, h^{-1}$ cMpc with a slope $\gamma = 2.00^{+0.14}_{-0.17}$ for our sample. The parameter constraints are shown in Figure 11. The best-fit parameters and the associated uncertainties (68% confidence intervals) are listed in Table 3.

### 5.2. Biased Dark Matter Model

The galaxy distribution is biased, relative to the underlying matter distribution, because galaxies form at peaks of the dark matter density fluctuations. A correlation function of galaxies is related to that of dark matter $\xi_{dm}(r)$ by the galaxy bias $b$ as,

$$\xi(r) = b^2 \xi_{dm}(r), \tag{17}$$

where we calculate $\xi_{dm}$ from the nonlinear matter power spectrum derived by Smith et al. (2003). For a scale-independent bias, Equation (17) can be simply rewritten for the projected form as

$$w_p(r_p) = b^2 w_{p,dm}(r_p), \tag{18}$$

where $w_{p,dm}$ is the projection of $\xi_{dm}$, calculated with $\pi_{max} = 30 \, h^{-1}$ cMpc, taking into account the Kaiser (1987) effect (see van den Bosch et al. 2013). We fit the $b^2 w_{p,dm}(r_p)$ to the 7–11th data points at $r_p > 1 \, h^{-1}$ cMpc, to avoid an enhancement of the resulting galaxy bias due to the significant

one-halo term of the data. We find $b = 2.44^{+0.38}_{-0.32}$ (see Table 3). The best-fit model and parameter constraints are shown in Figures 10 and 12, respectively.

### 5.3. Comparisons of Clustering Strength

In Figure 13, we compare the observed correlation length ($r_0 = 5.26^{+0.75}_{-0.62} \, h^{-1}$ cMpc) to other previous measurements, up to $z \sim 5$, and predictions from the dark halo model. Most past studies presented here used samples of star-forming galaxies with stellar masses or luminosity similar to our sample (typically $M_* \sim 10^{10} \, M_\odot$ and $L_{H\alpha} \sim 10^{42}$ erg s$^{-1}$), and thus they can be straightforwardly compared to each other.

At lower redshifts, Shioya et al. (2008) and Nakajima et al. (2008) measured clustering of narrow-band-selected H$\alpha$-emitters (HAEs) at $z \sim 0.24$ and $z \sim 0.4$, respectively. They found correlation lengths of $r_0 = 1.3 \, h^{-1}$ cMpc ($z \sim 0.24$) and $r_0 = 1.1 \, h^{-1}$ cMpc ($z \sim 0.4$), which are smaller than our results and others (filled triangles in Figure 13). Nakajima et al. (2008) mentioned that the low H$\alpha$ luminosity limits of their samples ($L_{H\alpha} \geqslant 10^{39.8}$ erg s$^{-1}$; roughly two orders of magnitude lower than that of our sample) and relatively large H$\alpha$ equivalent widths (i.e., their lower stellar masses), are likely responsible for the observed weak clustering strengths.

The HiZELS team conducted a wide-field, near-infrared narrow-band survey of HAEs. Sobral et al. (2010) presented a clustering analysis of HAEs at $z \sim 0.84$ and found $r_0 \sim 2.7 \, h^{-1}$ cMpc and $\sim 4.8 \, h^{-1}$ cMpc for their entire sample and a brighter subsample ($L_{H\alpha} \geqslant 10^{42}$ erg s$^{-1}$), respectively. The latter measurement is in rough agreement with our result, as expected from the similar H$\alpha$ luminosity limit. In addition, Geach et al. (2012) measured a correlation function at $z \sim 2.23$ with SFR $\gtrsim 7 \, M_\odot$ yr$^{-1}$, and found $r_0 \sim 3.7 \, h^{-1}$ cMpc. This slightly smaller correlation length (than ours and the HiZELS result at $z = 0.84$) may be reasonable, given the lower SFR limit for their sample.

There have been many studies that carry out a clustering analysis for color-selected star-forming galaxies. Adelberger et al. (2005) measured the clustering for BM ($z \sim 1.7$) and BX ($z \sim 2.2$) galaxies (Adelberger et al. 2004) with $r_0 = 4.5 \, h^{-1}$ cMpc and $4.2 \, h^{-1}$ cMpc. Hartley et al. (2010) presented a measurement for rest-frame $U - V$ selected star-forming galaxies and found that the correlation length of subsamples at a fixed $K$-band luminosity (a proxy of stellar mass) decreases from $z \sim 3$ to the present day. Their measurements at $z \sim 1.6$ are similar to ours. Bielby et al. (2014) used a sample of NUV $- r$ selected star-forming galaxies at $z \sim 0.5$–1.75, showing $r_0 \sim 4$–$5 \, h^{-1}$ cMpc for galaxies with log $M_*/M_\odot = 9.57$–11. Ishikawa et al. (2015) found $r_0 = 4.12 \pm 0.07 \, h^{-1}$ cMpc for $gzK$-selected star-forming galaxies with $K < 23$. The $BzK$ color selection is known to perform well at selecting star-forming (sBzK) galaxies at $1.4 \lesssim z \lesssim 2.5$ (Daddi et al. 2004). Hayashi et al. (2007) found $r_0 = 3.2 \, h^{-1}$ cMpc for a sample of sBzK galaxies with $K < 23.3$ and $\langle z \rangle = 1.9$. Hartley et al. (2008) measured a stronger clustering strength with $r_0 \sim 6.8 \, h^{-1}$ cMpc; they argued that contamination from highly clustered galaxies at higher redshifts ($z \gtrsim 2.5$) may be responsible for such a high amplitude. Lin et al. (2012) found $r_0 \sim 6.3 \, h^{-1}$ cMpc for sBzK galaxies of $10 < \log(M_*/M_\odot) < 10.5$. At $z \sim 1.6$–1.7, McCracken et al. (2010) measured an angular correlation function of $K_s < 23$ (22) sBzK galaxies in the COSMOS field with $r_0 \sim 4.25$ (4.69) $h^{-1}$ cMpc. Another study in COSMOS
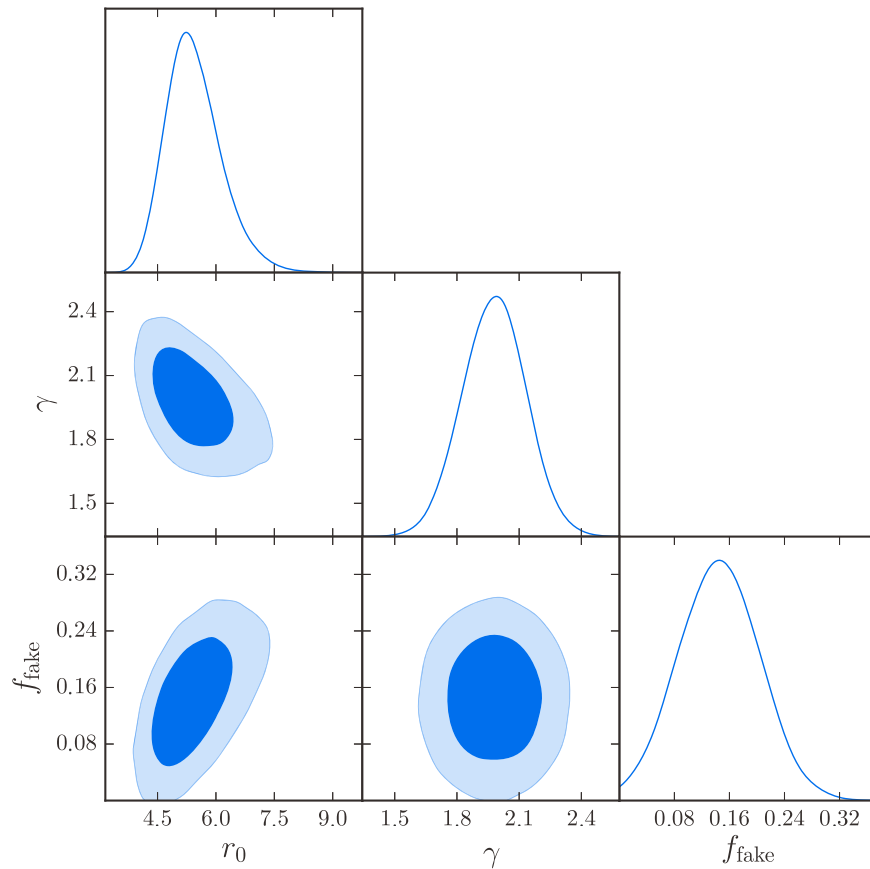
**Figure 11.** Constraints on the power-law parameters $(r_0, \gamma)$ and $f_{\rm fake}$. Contours show the 68 and 95% confidence levels. Solid lines show the posterior probability distribution of each parameter.

(Béthermin et al. 2014) found $r_0 = 4.0 \pm 0.9$ for sBzK galaxies of $M_* \sim 10^{10} M_\odot$. Although these measurements are slightly different from one one study to another, our measurement is in rough agreement and bracketed by measurements from past studies using a sample with a range of stellar mass or luminosity similar to our sample.

At higher redshifts, there have been efforts to measure the clustering of Lyman break galaxies (LBGs) at $z \sim 3$–5 (e.g., Foucaud et al. 2003; Adelberger et al. 2005; Kashikawa et al. 2006; Bielby et al. 2013) or beyond (Harikane et al. 2016), finding correlation lengths to be $r_0 \sim 4$–$5\,h^{-1}$ cMpc. Durkalec et al. (2015b) measured the projected correlation function of $\sim$3000 galaxies from the VIMOS Ultra Deep Survey, including the COSMOS field, and found $r_0 \sim 4\,h^{-1}$ cMpc at $z \sim 2.5$–5. These measurements are similar to those for star-forming galaxies at $z \sim 1$–2, including our own, although the average properties of LBGs may not be identical to lower-redshift color- or H$\alpha$-selected star-forming galaxies.

Given these comparisons, we conclude that our measurement is in general agreement with other previous measurements for star-forming galaxies at similar redshifts. In Figure 13, no clear evolutionary trend can be seen since $z \lesssim 4$. This is consistent with the moderate evolution of $r_0$ for a single population of galaxies, as predicted by the $\Lambda$CDM framework (e.g., Kauffmann et al. 1999). However, it is generally difficult to compare the clustering strengths at different redshifts because measurements are based on different selection functions.

## 6. Connection Between Galaxies and Dark Matter Halos

Knowledge of the connection between galaxies and dark matter halos is essential to understanding how and in what environments galaxies form and evolve. We investigate the properties of halos of "main-sequence" star-forming galaxies with $M_* \gtrsim 10^{9.57} M_\odot$ at $z \sim 1.6$ by interpreting the observed correlation function with an HOD model.

### 6.1. The Halo Model and Occupation Distribution

In the standard CDM paradigm, dark matter halos form at peaks in the matter density field. The global properties of halos, such as their abundance (or the halo mass function) and large-scale clustering amplitude (or the halo bias), are primarily determined by halo mass (Press & Schechter 1974; Cole & Kaiser 1989; Mo & White 2002). Galaxies reside in dark matter halos. Therefore, the observable abundance and clustering of galaxies can be used to constrain the connection between galaxies and dark matter halos.

The HOD framework is a convenient parametric way to describe the galaxy–dark matter connection, in order to model the abundance and clustering of galaxies. In its simplest form, the HOD describes the average number of galaxies $\langle N|M, z \rangle$ that reside in a halo of mass $M$ at redshift $z$, assuming that this number does not depend upon the formation history and environment of halos. For the model implemented here, we subdivide galaxies into being either central or satellite galaxies (e.g., Zheng et al. 2005), depending upon their location within

**Table 3**
Parameter Constraints

| Model | Params. | Prior[a] | Best-fit[b] | Best-fit ($f_{\text{fake}} = 0$)[c] |
|-------|---------|----------|-------------|-------------------------------------|
| Power-law | $r_0/(h^{-1}\,\text{cMpc})$ | ... | $5.26^{+0.75}_{-0.62}$ | $4.52^{+0.55}_{-0.57}$ |
| | $\gamma$ | ... | $2.00^{+0.14}_{-0.17}$ | $2.00^{+0.16}_{-0.20}$ |
| | $f_{\text{fake}}$ | $\geqslant 0$, $G(0.14, 0.06)$ | $0.14^{+0.06}_{-0.06}$ | ... |
| Dark matter | $b$ | ... | $2.44^{+0.38}_{-0.32}$ | $2.10^{+0.33}_{-0.32}$ |
| | $f_{\text{fake}}$ | $\geqslant 0$, $G(0.14, 0.06)$ | $0.14^{+0.06}_{-0.06}$ | ... |

**Notes.**
[a] Prior probability distribution for each parameter. $G(x_0, \sigma)$ denotes a Gaussian function with a mean $x_0$ and standard deviation $\sigma$.
[b] The best-fitting model parameters that give the minimum $\chi^2$ with the broad prior on $f_{\text{fake}}$ (see Section 4.3).
[c] The best-fitting parameters for $f_{\text{fake}} = 0$.
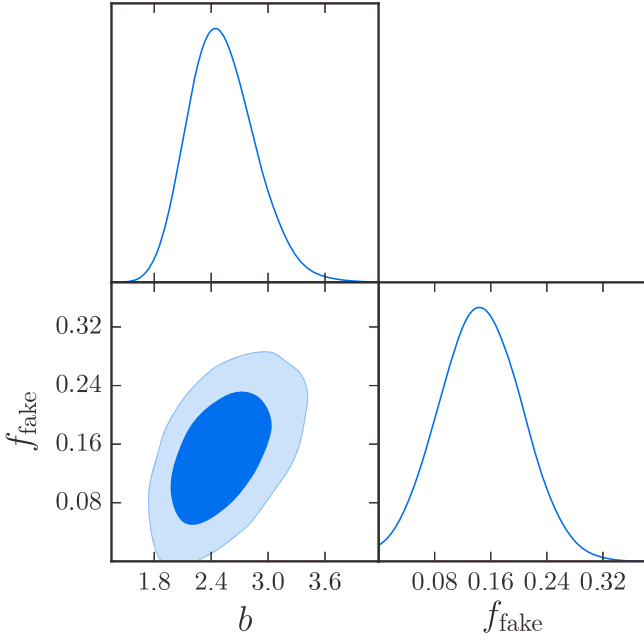


**Figure 12.** Constraints on the galaxy bias $b$ and $f_{\text{fake}}$. Contours show the 68 and 95% confidence levels. Solid lines show the posterior probability distribution of each parameter.

their dark matter halos, such that

$$\langle N|M, z\rangle = \langle N_{\text{cen}}|M, z\rangle + \langle N_{\text{sat}}|M, z\rangle. \quad (19)$$

The average number density of galaxies is then simply given by

$$n_{\text{tot}} = \int \langle N|M, z\rangle n(M, z)\,dM, \quad (20)$$

where the halo mass function $n(M, z)\,dM$ gives the number density of halos of mass $M \pm dM/2$ at redshift $z$. The average number density of central ($n_{\text{cen}}$) or satellite galaxies ($n_{\text{sat}}$) is calculated by replacing $\langle N|M, z\rangle$ in Equation (20) with $\langle N_{\text{cen}}|M, z\rangle$ or $\langle N_{\text{sat}}|M, z\rangle$. For calculations, we fix the redshift to the median of the sample ($z = 1.588$). Hereafter, the variable $z$ is omitted from equations.

The clustering of galaxies can be quantified as the excess probability over the random case of finding two galaxies separated by a distance. It is described by the two-point correlation function $\xi$. The clustering of galaxies arises from a combination of two contributions. The one-halo term

corresponds to the pairs of galaxies that reside within the same halo, and the two-halo term to the pairs of galaxies that reside in distinct halos:

$$\xi(r) = \xi^{1h}(r) + \xi^{2h}(r), \quad (21)$$

where the superscripts "1h" and "2h" stand for the one-halo and the two-halo terms, respectively. The correlation function $\xi(r)$ and the power spectrum $P(k)$ form a Fourier transform pair such that

$$\xi(r) = \frac{1}{2\pi^2} \int_0^\infty dk\, k^2 P(k) \frac{\sin kr}{kr}, \quad (22)$$

which implies that Equation (21) can also be written as

$$P(k) = P^{1h}(k) + P^{2h}(k). \quad (23)$$

The one-halo term can be further expressed as the sum of contributions from the central-satellite and satellite-satellite galaxy pairs hosted by the same halo:

$$P^{1h}(k) = \frac{1}{n_{\text{tot}}^2} \int n(M)\,dM \quad (24)$$

$$[2\langle N_{\text{cen}}|M\rangle \langle N_{\text{sat}}|M\rangle u(k, M) + \langle N_{\text{sat}}|M\rangle^2 u^2(k, M)], \quad (25)$$

where $u(k, M)$ describes the Fourier transform of the density profile of satellite galaxies within dark matter halos. We assume that central galaxies reside at the center of halos and the occupation numbers of centrals and satellites are independent of each other, such that $\langle N_{\text{cen}} N_{\text{sat}}\rangle = \langle N_{\text{cen}}\rangle \langle N_{\text{sat}}\rangle$. We have further assumed that $N_{\text{sat}}$ follows Poisson statistics, such that $\langle N_{\text{sat}}(N_{\text{sat}} - 1)\rangle = \langle N_{\text{sat}}\rangle^2$. This is supported both by observations (Yang et al. 2008) and numerical simulations (Kravtsov et al. 2004; Zheng et al. 2005).

The two-halo term consists of contributions from the central–central, central–satellite, and satellite–satellite galaxy pairs hosted by distinct halos:

$$P^{2h}(k) = \frac{1}{n_{\text{tot}}^2} \int dM_1 \int dM_2\, n(M_1)n(M_2) \quad (26)$$

$$[\langle N_{\text{cen}}|M_2\rangle \langle N_{\text{cen}}|M_2\rangle + 2\langle N_{\text{cen}}|M_1\rangle \langle N_{\text{sat}}|M_2\rangle u(k, M_2) \quad (27)$$

$$+ \langle N_{\text{sat}}|M_1\rangle \langle N_{\text{sat}}|M_2\rangle u(k, M_1)u(k, M_2)]P_{\text{hh}}(k|M_1, M_2), \quad (28)$$

where $P_{\text{hh}}(k|M_1, M_2)$ describes the cross-power spectrum of halos of masses $M_1$ and $M_2$. Following van den Bosch et al. (2013), we express it as a product of the large-scale bias of halos of masses $M_1$ and $M_2$, and the nonlinear matter power
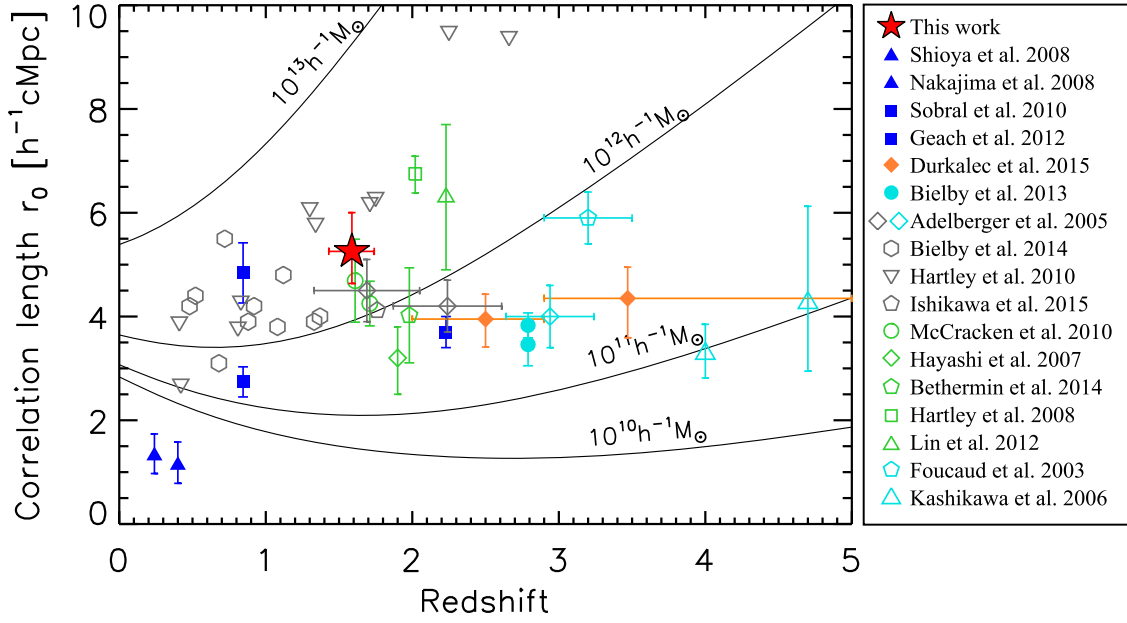
**Figure 13.** Correlation length $r_0$ as a function of redshift. The red star indicates our measurement. Open and filled symbols show measurements from the literature based on photometric and spectroscopic (or narrow-band imaging) observations, respectively. Blue: H$\alpha$ emitters in narrow-band surveys. Orange: VIMOS Ultra Deep Survey. Green: sBzK galaxies. Gray: other color-selected star-forming galaxies. Cyan: LBGs. Four solid curves indicate the correlation length of dark halos of different masses, as labeled.

spectrum (Smith et al. 2003), and account for the radial dependence of the bias as well as halo exclusion.

We assume that the HOD of central galaxies is given by

$$\langle N_{\rm cen}|M\rangle = \frac{1}{2}\left[1 + {\rm erf}\left(\frac{\log M - \log M_{\rm min}}{\sigma_{\log M}}\right)\right], \qquad (29)$$

and that of satellite galaxies is given by

$$\langle N_{\rm sat}|M\rangle = \frac{1}{2}\left[1 + {\rm erf}\left(\frac{\log(M/M_{\rm min})}{\sigma_{\log M}}\right)\right]\left(\frac{M - M_{\rm cut}}{M_1'}\right)^{\alpha}. \quad (30)$$

For $M < M_{\rm cut}$, $\langle N_{\rm sat}|M\rangle = 0$. The parameter $M_{\rm min}$ is a halo mass above which a halo has a central galaxy. This transition is relaxed with a smoothing scale $\sigma_{\log M}$ (Zheng et al. 2005, 2007). The average number of satellite galaxies increases with increasing halo mass by a power law parameterized by a slope $\alpha$ and normalization $M_1'$, which is related to the halo mass ($M_1$) at which a halo is expected to have a single satellite galaxy, as $M_1 = M_1' + M_{\rm cut}$.

For computing all observables, we use the halo mass function from Tinker et al. (2010), which defines a halo as a spherically collapsed region with an average density 200 times greater than the background matter density of the universe, and a large-scale halo bias proposed by Tinker et al. (2010), with an empirical radial scale dependence derived by Tinker et al. (2012) and corrections described in van den Bosch et al. (2013). We also assume that the radial distribution of satellite galaxies follows the density distribution of dark matter in halos. For this purpose, we use the Navarro–Frenk–White profile (Navarro et al. 1997), with the mass-concentration relation calibrated by Macciò et al. (2007). In addition, we take into account the prescriptions of Kaiser (1987) effect to infer $w_{\rm p}(r_{\rm p})$, following the recipe described in van den Bosch et al. (2013).

Once a set of HOD parameters is given, the following physical quantities are inferred:

1. Effective large-scale bias

$$b_{\rm eff} = \frac{1}{n_{\rm tot}}\int b_h(M)\langle N|M\rangle n(M)dM \qquad (31)$$

2. Satellite fraction

$$f_{\rm sat} = \frac{1}{n_{\rm tot}}\int \langle N_{\rm sat}|M\rangle n(M)dM \qquad (32)$$

3. Effective halo mass

$$M_{\rm eff} = \frac{1}{n_{\rm tot}}\int M\langle N|M\rangle n(M)dM. \qquad (33)$$

The effective large-scale bias is the number-weighted average of the halo bias $b_h(M)$. The effective halo mass is the number-weighted average mass of halos that host galaxies in the sample.

### 6.2. Limitations for the HOD Parameters

We sample the posterior distribution of the HOD parameters, given the abundance and clustering measurements, using a MCMC technique. Although the HOD model defined above has five parameters, our data do not have enough statistics to simultaneously constrain all the parameters. We here describe the prior limitations imposed on some model parameters to resolve degeneracies and avoid overfitting.

We fix the power-law slope $\alpha$ in Equation (30) to be 1, which is supported observationally (Zehavi et al. 2005; Zheng et al. 2007) and theoretically (Kravtsov et al. 2004; Zheng et al. 2005), and has been commonly applied in past studies (e.g., Conroy et al. 2006). We further impose a relation between $M_{\rm cut}$

and $M_1'$ as

$$\log M_{\rm cut}/(h^{-1}M_\odot) = 0.76 \log M_1'/(h^{-1}M_\odot) + 2.3, \quad (34)$$

following Conroy et al. (2006). In addition, we independently impose priors on $\sigma_{\log M}$ by using a stellar-to-halo mass relation derived by Behroozi et al. (2013a), with the uncertainties on the stellar mass estimate taken into account (see Appendix C for details). The prior probability distribution is given by

$$P(\sigma_{\log M}) \propto \begin{cases} \exp\left[-\frac{(\sigma_{\log M}-0.24)^2}{2\times 0.03^2}\right] & \text{for } \sigma_{\log M} \geqslant 0, \\ 0 & \text{for } \sigma_{\log M} < 0. \end{cases} \quad (35)$$

The prior on the contamination fraction $f_{\rm fake}$ is given as $f_{\rm fake} = 0.14 \pm -0.06$ (see Section 4.3). We use uniform non-informative priors on the halo mass parameters $\log M_{\rm min}/(h^{-1}M_\odot)$ and $\log M_1'/(h^{-1}M_\odot)$, in the range [9, 15].

With the HOD model parameterized by Equations (29)–(30), the number density of all galaxies with $M_* \geqslant M_*^{\rm lim}$, i.e., the $M_*$-selected sample, can be predicted using Equation (20), then compared to the observed number density of $9.29 \times 10^{-3}\,h^{-3}\,{\rm cMpc}^{-3}$ (see Table 2). We emphasize that this is not the number density of the FMOS-parent sample with additional selection based on $K_S$ and predicted H$\alpha$ flux. In doing so, we have implicitly assumed that our FMOS-spec-$z$ sample is representative of the $M_*$-selected sample (see Section 2.3). We estimate the cosmic variance of the number density by using the subhalo mock catalog from $\nu^2$GC simulation, and find approximately 10% fluctuation from one to another sub-box with the same volume as our survey. Therefore, we use the observed abundance to be $n_{\rm tot} = (9.29 \pm 0.93) \times 10^{-3}\,h^{-3}\,{\rm cMpc}^{-3}$ as an additional independent observable to be compared to the predictions in the MCMC procedure. The prior information is summarized in Table 3.

### 6.3. HOD Model Fit

We compare the projected correlation function computed from the HOD model parameters to the observed $w_{\rm p}(r_{\rm p})$. We have varied three HOD parameters ($\log M_{\rm min}$, $\sigma_{\log M}^2$, $\log M_1'$) plus the contamination fraction $f_{\rm fake}$, while using 11 data points at $-1.5 < \log(r_{\rm p}/h^{-1}\,{\rm cMpc}) < 1.1$ and an independent observational constraint on the galaxy number density $n_{\rm tot}$ (see Section 6.2). As a result, we have eight degrees of freedom in this analysis (or nine for the case with a fixed $f_{\rm fake} = 0$). The three HOD parameters can be effectively constrained using our observables. The posterior distribution of the HOD and the parameters are shown in Figures 14 and 15, respectively. In addition, we have also calculated various physical quantities using Equations (20) and (31)–(33). The best-fit parameters and the inferred quantities are summarized in Table 4. The posterior probability distributions are shown in Figure 16. It can been seen that $M_{\rm min}$ is degenerate with $M_1'$. This negative correlation is mainly caused by the prior constraint on the total abundance of galaxies, because a smaller $M_{\rm min}$ leads to a more abundant number of centrals; thus, $M_1'$ needs to increase to reduce the number of satellites, and vice versa. This would reduce the amplitude of clustering, which is compensated by a smaller $f_{\rm fake}$, as reflected by the positive (negative) correlation between $f_{\rm fake}$ and $M_{\rm min}$ ($M_1'$) in Figure 15. It is worth noting that the data prefers a lower value for the fraction of fake detections $f_{\rm fake}$ compared to the prior (shown by dotted line in Figure 15). We
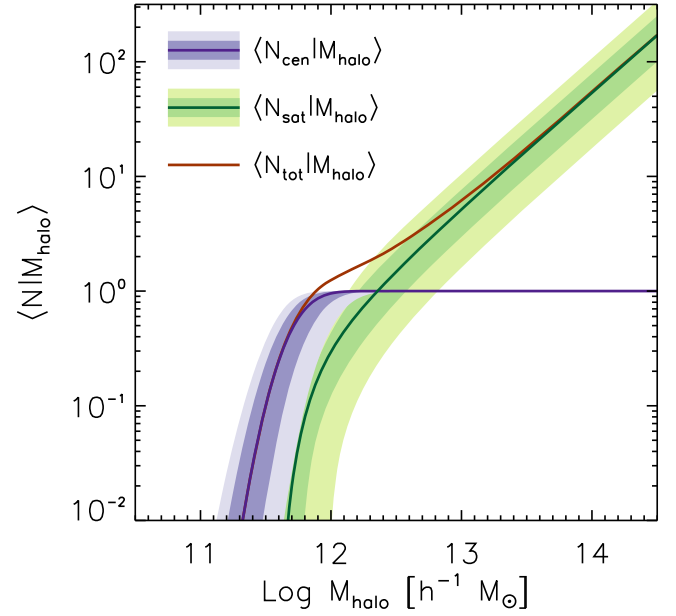


**Figure 14.** Halo occupation distribution for our sample. Purple, green, and dark red lines show the average numbers of central, satellite, and all galaxies in a halo as a function of halo mass, respectively. Dark and light shaded regions show the 68 and 95% confidence intervals, respectively.

note that the results for $f_{\rm fake} = 0$ are in agreement with the fiducial analysis within $1\sigma$ (the rightmost column in Table 4). For the remainder of the paper, we use the results from the fiducial prior on $f_{\rm fake}$.

In Figure 17, we show the observed $w_{\rm p}(r_{\rm p})$ and the model computed from the best-fitting HOD parameters. The observed $w_{\rm p}(r_{\rm p})$ is corrected for the effect of fake detections with the best-fit value of $f_{\rm fake} = 0.099$. The data points are well-fit with the model, with a clear signature of the one-halo term. We find that the transition from the one-halo-dominated regime at small scales to the two-halo-dominated regime at large scales happens at $r_{\rm p} \approx 0.5\,h^{-1}\,{\rm cMpc}$. Because the two-halo term dominates at scales larger than the typical virial radius of largest halos, this scale is a useful probe of the virial size of halos that host the galaxies. For SDSS galaxies ($z \lesssim 0.1$), such a transition happens around $r_{\rm p} = 1$–$2\,h^{-1}\,{\rm cMpc}$ (Zehavi et al. 2004). In contrast, Zheng et al. (2007) measured the one-to-two halo transition at $r_{\rm p} \sim 0.4$–$0.6\,h^{-1}\,{\rm cMpc}$ at $z \sim 1$, which is in good agreement with our result. In addition, Gobat et al. (2015) found an excess of the number of satellite galaxies up to $\sim 300$ physical kpc away from the halo center at a similar redshifts ($z \sim 1.8$) through image stacking. This putative halo radius, based on an independent technique, is fully consistent with our measurement. These results show that the transition scale decreases for high redshift galaxies, and this trend likely reflects the fact that fewer massive clusters have formed at earlier epochs.

## 7. Discussion

In this section, we discuss the properties of halos that contain star-forming galaxies at $z \sim 1.6$ based on the physical parameters inferred from the HOD modeling of the observed correlation function, and then present a new constraint on the stellar-to-halo mass relation. At the end, we demonstrate the capabilities of a future survey with a next-generation multi-object spectrograph.
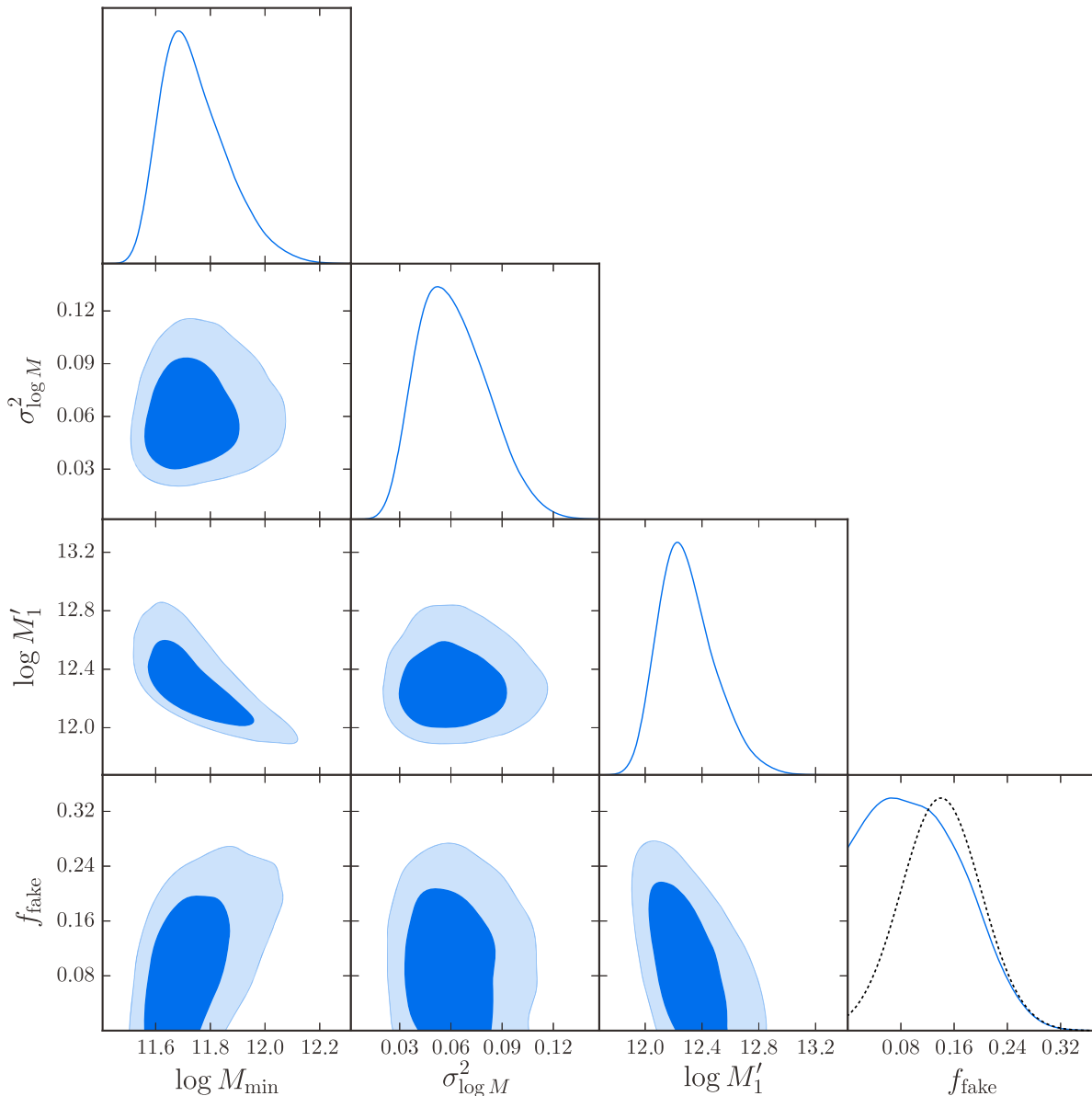
**Figure 15.** Constraints of the HOD parameters ($\log M_{\min}$, $\sigma_{\log M}^2$, $\log M_1'$) and $f_{\text{fake}}$. Contours show the 68 and 95% confidence levels. Solid lines show the posterior probability distribution of each parameter. A dashed line indicates the prior probability distribution of $f_{\text{fake}}$.

### 7.1. Halo Mass and Large-scale Bias

We find the effective halo mass to be $M_{\text{eff}} = 4.71_{-1.62}^{+1.19} \times 10^{12}\, h^{-1}\, M_\odot$ (Equation (33)), and the effective large-scale galaxy bias to be $b_{\text{eff}} = 2.18_{-0.24}^{+0.18}$ (Equation (31)). We note that this bias estimation is fully consistent with what was found for the biased dark matter fitting in Section 5.2, while taking into account the difference in the best-fit values of $f_{\text{fake}}$.

In Figure 18, we present $M_{\text{eff}}$ in comparison with the average growth histories of halos with different present-day masses, as derived by Zhao et al. (2009) and Behroozi et al. (2013a). We find that $M_{\text{eff}}$ of our sample lies on the mass-assembly history of halos with a present-day mass $M_{\text{h}}$ ($z = 0$) $\approx 2 \times 10^{13}\, h^{-1}\, M_\odot$ (thick gray curve). This is equivalent to the typical mass of group-scale halos. We also plot effective halo masses, derived through HOD modeling of galaxy clustering at different redshifts, from the literature. At lower redshifts, we present measurements from the CFHT Legacy Survey at

$z \sim 0.3$, 0.5, and 0.7 (Coupon et al. 2012). These samples are selected based on the absolute g-band magnitude ($M_g - 5 \log h < -19.8$). The galaxy number densities are $n_{\text{gal}} \approx 8 \times 10^{-3}\, h^3\,\text{cMpc}^{-3}$, similar to our $M_*$-selected sample. In addition, the results from Abbas et al. (2010) at $0.2 \lesssim z \lesssim 1.3$ based on the VIMOS-VLT Deep Survey (VVDS; Le Fèvre et al. 2005), are shown. We highlight two subsamples in Abbas et al. (2010) with $M_B < -19.5$ at $z = 0.67$ and 0.99 (filled triangles) that have a number density similar to our $M_*$-selected sample. We also present the results for samples with $M_* > 10^{10}\, M_\odot$ $z \sim 1.1$ and $\sim 1.5$ from the NEWFIRM Medium Band Survey (Wake et al. 2011). These samples at lower redshifts are in excellent agreement with the same halo mass assembly history as our data. At higher redshifts, we show results from HiZELS ($z \sim 2.2$; Geach et al. 2012) and VUDS ($z \sim 2.5$ and $\sim 3.5$; Durkalec et al. 2015b), which are in broad agreement as well, favoring a slightly lower

**Table 4**
Priors and Constraints of the HOD Parameters

| Parameters | Prior[a] | Best-fit | Best-fit ($f_{fake}$=0) |
|---|---|---|---|
| $\log M_{min}/(h^{-1}M_\odot)$ | $U(9, 15)$ | $11.72^{+0.11}_{-0.12}$ | $11.65^{+0.10}_{-0.10}$ |
| $\sigma^2_{\log M}$ | $\geqslant 0$, $G(0.24, 0.03)$ for $\sigma_{\log M}$ | $0.057^{+0.021}_{-0.020}$ | $0.058^{+0.023}_{-0.024}$ |
| $\log M'_1/(h^{-1}M_\odot)$ | $U(9, 15)$ | $12.27^{+0.17}_{-0.21}$ | $12.39^{+0.21}_{-0.24}$ |
| $f_{fake}$ | $\geqslant 0$, $G(0.14, 0.06)$ | $0.098^{+0.050}_{-0.091}$ | ... |
| $\chi^2/\nu$[b] | ... | 1.45 ($\nu = 8$) | 1.10 ($\nu = 9$) |

| Inferred quantities | Constraint | Best-fit | Best-fit ($f_{fake} = 0$) |
|---|---|---|---|
| $n_{tot}/(h^{-1}cMpc)^3$[c] | $>0$, $G(9.29\,0.93)$ | $9.22^{+1.33}_{-1.32} \times 10^{-3}$ | $9.17^{+1.52}_{-1.43} \times 10^{-3}$ |
| $n_{cen}/(h^{-1}\,cMpc)^3$ | ... | $5.61^{+1.52}_{-1.73} \times 10^{-3}$ | $6.60^{+1.58}_{-1.68} \times 10^{-3}$ |
| $n_{sat}/(h^{-1}\,cMpc)^3$ | ... | $3.61^{+1.39}_{-2.13} \times 10^{-3}$ | $2.57^{+1.00}_{-1.97} \times 10^{-3}$ |
| $f_{sat}$ | ... | $0.39^{+0.15}_{-0.20}$ | $0.28^{+0.11}_{-0.20}$ |
| $M_{eff}/(h^{-1}M_\odot)$ | ... | $4.71^{+1.19}_{-1.62} \times 10^{12}$ | $3.85^{+0.97}_{-1.49} \times 10^{12}$ |
| $b_{eff}$ | ... | $2.18^{+0.18}_{-0.24}$ | $2.05^{+0.15}_{-0.22}$ |
| $\log M_1/M_{min}$ | ... | $0.64^{+0.30}_{-0.28}$ | $0.82^{+0.30}_{-0.29}$ |

**Notes.**
[a] Prior probability distribution of each parameter. Here, $U(x_1, x_2)$ is a non-informative prior with a interval $[x_1, x_2]$. $G(x_0, \sigma)$ is a Gaussian function with a mean $x_0$ and standard deviation $\sigma$.
[b] $\chi^2$ counts the contributions from 11 data points of $w_p(r_p)$ plus the galaxy number density $n_{tot}$.
[c] Observed number density of the $M_*$-selected sample. This is independently constrained from the photometric catalog.

mass ($M_h \approx 10^{13} h^{-1} M_\odot$ at $z = 0$). In summary, these results indicate a good agreement between the predictions of the halo mass assembly history and observations over a wide range of cosmic history, from $z \sim 4$ to the present. From another perspective, there is evidence that these samples at different redshifts essentially represent similar galaxy populations at different epochs.

In Figure 19, we show the effective large-scale galaxy bias, compared to measurements from the literature that are shown in Figure 18. Our data are in broad agreement with measurements at $1 \lesssim z \lesssim 2$ by Geach et al. (2012) and Wake et al. (2011). Based on measurements over a wide range of redshift, it seems that the galaxy bias decreases with cosmic time. This trend agrees with the general expectation from the standard scenario of hierarchical structure formation (e.g., Mo & White 1996; Kauffmann et al. 1999). For reference, we show the halo bias for fixed halo masses ($\log M_h/(h^{-1}M_\odot) = 10, 11, 12, 13,$ and 14; thin solid lines), which present a rapid decline of the bias. As mentioned above, halos of a present-day mass $M_h(z = 0) = 2 \times 10^{13} h^{-1} M_\odot$ are expected to be the descendants of halos containing star-forming galaxies in our sample at $z \sim 1.6$ (thick gray line in Figure 18). We further calculate the halo bias for the evolving halo mass as a function of redshift. The global trend of the effective bias is well-represented by this evolutionary track (thick gray line in Figure 19).

### 7.2. Ratio of Stellar Mass to Halo Mass

The ratio of stellar mass to halo mass ($M_*/M_h$) encodes the efficiency of converting baryons into stars, relative to the total amount of dark matter accreting onto halos. Yang et al. (2003) proposed a functional form with double power-law components to express the average ratios of halo mass to galactic luminosity, as a function of halo mass, motivated by the fact that the observed luminosity function is steeper (shallower) than the halo mass function at the high (low) mass end (see also Vale & Ostriker 2004; Cacciato et al. 2013). With improvement of stellar mass measurements, this formalism has been applied

to describe the stellar-to-halo mass relation (SHMR; $M_*/M_h$ versus $M_h$) (e.g., Zheng et al. 2007; Yang et al. 2008, 2009). Past studies based on large data sets (e.g., SDSS) have shown that the $M_*/M_h$ ratio reaches a peak around a halo mass of $M_h \sim 10^{12} M_\odot$ at $z \sim 0$–1 (e.g., Behroozi et al. 2010; Guo et al. 2010; Moster et al. 2010, 2013; More et al. 2011; Leauthaud et al. 2012; Coupon et al. 2015). Beyond $z > 1$, there have been limited efforts to measure the SHMR based on HOD modeling of galaxy clustering (Wake et al. 2011; Geach et al. 2012; Durkalec et al. 2015a, 2015b; Martinez-Manso et al. 2015; Harikane et al. 2016). Other studies use an alternative technique (i.e., abundance matching; Kravtsov et al. 2004) to predict the SHMRs (Behroozi et al. 2013a; Moster et al. 2013). These studies model a numerical form of the SHMR and use it to populate halos in a $N$-body simulation with galaxies, then the model is adjusted to match the inferred galaxy abundance to observations.

For our sample of star-forming galaxies, we find $\log M_*/M_h = -2.3 \pm 0.2$ at $M_h = M_{min} = 10^{11.87} M_\odot$ (computed with $h = 0.7$), which is determined as the ratio of the threshold stellar mass $M_*^{lim}$ to $M_{min}$. Figure 20 presents the observed $M_*/M_h$ ratio as a function of $M_h$ (red star). Red and blue contours indicate the 68 and 95% confidence intervals for the cases with a non-zero $f_{fake}$ and $f_{fake} = 0$, respectively. We emphasize that our result probes the low-mass side of the SHMR at this epoch, based on a spectroscopic sample, evidently confirming the rapid decline at $M_h < 10^{12} M_\odot$. We compare our result to past measurements at similar redshifts from the literature, which are based on HOD modeling of galaxy clustering via photometric redshifts and/or color-selected galaxy samples. The values of stellar mass shown here are derived with (or converted to) a Chabrier IMF and a Hubble parameter $h = 0.7$. Martinez-Manso et al. (2015) measure the $M_*/M_h$ ratios over $12 \lesssim \log M_h/M_\odot \lesssim 13.2$ by using a sample of $z \sim 1.5$ galaxies selected from the *Spitzer*/IRAC 3.6 and 4.5 $\mu$m photometry. Their result shows an evident peak at $M_h = 10^{12.4} M_\odot$. Around this peak halo mass,
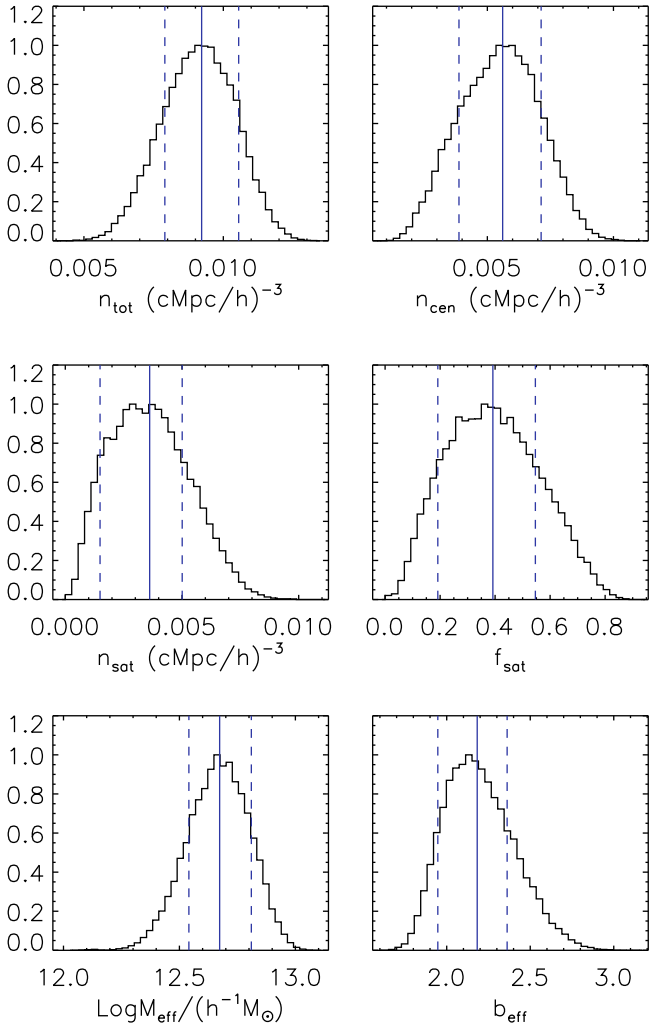
**Figure 16.** Posterior distribution of the inferred physical quantities. From the top to bottom: average total number density ($n_{tot}$), average number density of central ($n_{cen}$) and satellite galaxies ($n_{sat}$), satellite fraction ($f_{sat}$), effective halo mass ($M_{eff}$), and effective large scale bias ($b_{eff}$). Vertical solid and dashed lines indicate the best-fit value giving a minimum $\chi^2$ and the 68% confidence interval, respectively.



**Figure 17.** Observed projected correlation function (filled circles) and the model from the best-fitting HOD (thick solid line). Green and orange thin solid curves show the one- and two-halo terms, respectively. Dark and light shaded regions indicate the 68 and 95% confidence intervals. The data points are corrected for the effect of fake detections with the best-fit $f_{fake} = 0.099$.



**Figure 18.** The average host halo mass as a function of redshift. A solid star indicates the effective halo mass $M_{eff}$ estimated for our sample using HOD modeling. Other symbols indicate results from the literature: filled green circles (error bars are comparable to the size of the symbol)—Coupon et al. 2012; empty triangle—Abbas et al. 2010; filled purple triangles—$M_B < -19.5$ samples from Abbas et al. 2010; empty circles—Wake et al. 2011; filled blue square—Geach et al. 2012; filled orange diamonds—Durkalec et al. 2015b. Solid and dashed curves show the mass assembly histories of halos for different present-day masses, derived by Behroozi et al. (2013a) (solid lines) and Zhao et al. (2009) (dashed lines), respectively. A thick gray line highlights the history of the present-day group-scale halos with $M_h$ ($z = 0$) = $2 \times 10^{13} h^{-1} M_\odot$, which well-represents the global trend of data points shown here.

Wake et al. (2011) and McCracken et al. (2015) measure $M_*/M_h$ at $z \sim 1.5$, which are systematically higher than our result.

We note that the $M_*/M_h$ ratios in Wake et al. (2011) are multiplied by a factor of 1.5 to account for the fact that the Maraston (2005) stellar population synthesis model induces stellar masses that are systematically lower than those derived by the Bruzual & Charlot (2003) model (by an average of $\sim$60%; Maraston et al. 2006). The measurements by Wake et al. (2011) are in good agreement with the result of Martinez-Manso et al. (2015), whereas the measurements by McCracken et al. (2015) are systematically greater by a factor of $\sim$2 at fixed $M_h$. Another study (Gobat et al. 2015) compares the average stellar mass of a sample of massive star-forming galaxies at $z \sim 1.8$ ($\sim$1.3 $\times 10^{11} M_\odot$) and the halo mass ($\sim$2.4 $\times 10^{13} M_\odot$) induced from the average X-ray luminosity through stacking analysis. Their independent measurement clearly indicates, as do others, the decline of SHMR at high masses. It may be worth noting that our measurement is in good agreement with the extrapolation of the SHMR derived by Martinez-Manso et al. (2015) (thin dashed line in Figure 20).
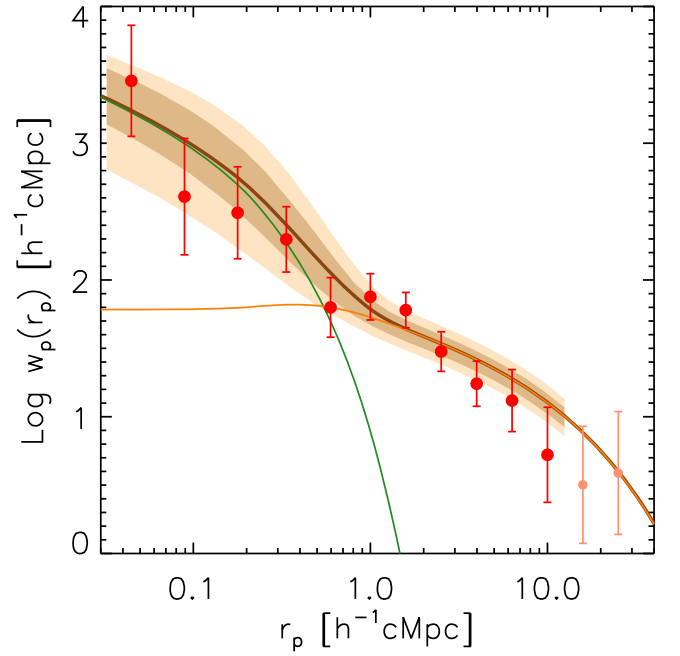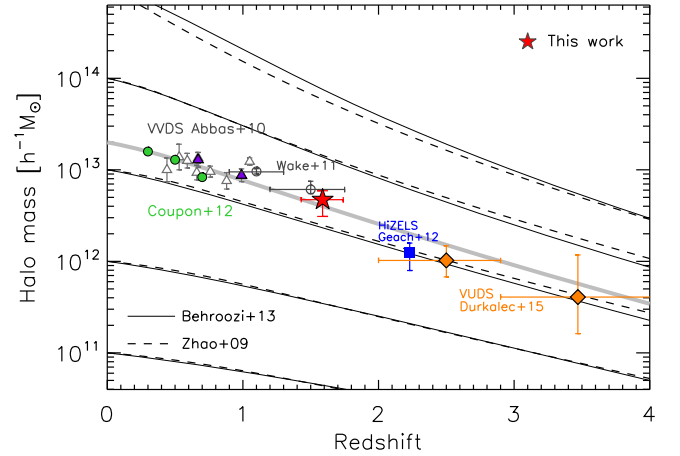
Furthermore, in Figure 20 we show the SHMRs at $z = 1.6$ that are derived by an abundance-matching technique (Behroozi et al. 2013a; Moster et al. 2013). The Behroozi et al. (2013a) SHMR presents a peak value that is higher (by $\sim$0.2 dex) than in Martinez-Manso et al. (2015), but more consistent with McCracken et al. (2015). Among these studies, the three
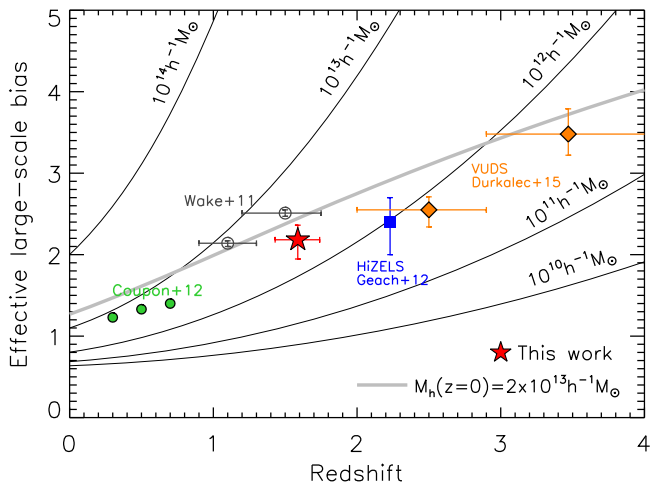
**Figure 19.** Galaxy bias as a function of redshift. The effective large-scale bias, based on HOD analysis, for our sample is shown by a red star. Other symbols indicate results at different redshifts in the literature: filled green circles— Coupon et al. 2012; empty circles—Wake et al. 2011; solid blue square— Geach et al. 2012; solid orange diamonds—Durkalec et al. 2015b. Solid curves indicate the relation among the bias, halo mass, and redshift, as labelled. The gray thick line indicates the evolutionary track of the bias of halos of a present-day mass $M_h(z=0) = 2 \times 10^{13} \, h^{-1} \, M_\odot$.
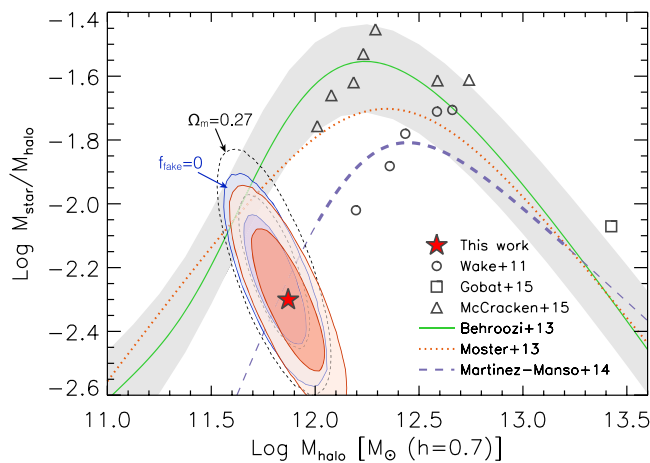


**Figure 20.** Ratio of stellar mass to halo mass ($M_*/M_h$), as a function of halo mass (SHMR). The star symbol indicates the ratio $M_*^{\rm lim}/M_{\rm min}$, with the 68 and 95% confidence intervals shown by red contours. The constraint for $f_{\rm fake} = 0$ is shown by blue contours. Dotted contours indicate the constraint with an alternative cosmology ($\Omega_m = 0.27$, see text). A dashed line indicates the best-fit relation derived by Martinez-Manso et al. (2015) at $12 < \log M_h < 13.2$ (thick dashed line) and its extrapolation (thin dashed line). Open circles show the measurements by Wake et al. (2011) for their different stellar mass threshold samples at $1.2 < z < 1.77$. Triangles and a square indicate results for samples at $z \sim 1.8$ (Gobat et al. 2015; McCracken et al. 2015). The SHMRs based on the abundance matching technique are shown as labeled (Behroozi et al. 2013a; Moster et al. 2013). A shaded region indicates the systematic uncertainties of the Behroozi et al. (2013a) relation.

measurements based on the HOD modeling (our result, Martinez-Manso et al. 2015, and Wake et al. 2011) likely favor an SHMR with a low-mass part that is systematically lower than that predicted by abundance matching. However, we note that our result is consistent with those at a $2\sigma$ confidence level, and becomes closer to the Behroozi et al. (2013a) SHMR (nearly $1\sigma$) when recalculated with the same cosmology as the reference ($\Omega_m = 0.27$; dotted contours in Figure 20). For quantitative derivation and comparison of SHMRs, one needs to carefully

treat the difference between adopted cosmologies (e.g., $\Omega_m$, $\sigma_8$) from one study to another. We refer the reader to More (2013) for further discussions of such systematic effects on the galaxy clustering.

### 7.3. Baryon Conversion Efficiency

Given an average growth history of halos and an observation of SFRs of galaxies, we can calculate the efficiency of baryon conversion $\epsilon_{\rm baryon}$, i.e., the fraction of the mass of baryons converted into stars per unit time to the total accretion rate of baryons falling into halos. For this purpose, the SFRs of our spec-$z$ galaxy sample are measured from the observed H$\alpha$ luminosities by using Equation (1). To derive intrinsic SFRs, we correct the observed H$\alpha$ flux for aperture loss (Silverman et al. 2015) and dust extinction, assuming a Calzetti et al. (2000) reddening curve. The level of extinction $E_{\rm star}(B - V)$ is estimated based on the $B_z - j$ color (Daddi et al. 2007), and converted to the attenuation toward nebular emission lines, assuming a relation $E_{\rm neb}(B - V) = E_{\rm star}(B - V)/0.66$ (see Kashino et al. 2013). We find the average SFR to be $\approx 15 \, M_\odot \, {\rm yr}^{-1}$ at $M_*^{\rm lim} = 10^{9.57} \, M_\odot$, with a standard deviation of 0.3 dex, including errors on the flux measurement. However, this average may be slightly biased toward a higher value than the $M_*$-selected sample because of the imposed limits on predicted H$\alpha$ flux and detection bias (see Section 2.2). According to SED-based SFRs shown in Figure 2, the median SFR in the spec-$z$ sample is elevated by approximately 0.1 dex (a factor of 1.3) as compared to the entire main sequence population ($K_S \leqslant 23.5$) at the threshold stellar mass. We do not take this bias into account because it is well below the scatter in SFR, and thus does not affect our estimate of the baryon conversion efficiency with the given statistics.

We calculate the mass accretion rates by using the halo mass accretion history derived by Behroozi et al. (2013a). With the baryon fraction of accreting matter fixed to the cosmic baryon fraction ($\Omega_b/\Omega_m = 0.17$), the baryon accretion rate is found to be $\dot{M}_{\rm baryon} = 44 \, M_\odot \, {\rm yr}^{-1}$ for halos of $M_{\rm halo} = M_{\rm min} = 10^{11.87} \, M_\odot$. We note that the value of $\dot{M}_{\rm baryon}$ varies almost proportionally with halo mass. As a result, we find that the main sequence galaxies with $M_* \sim 10^{9.57} \, M_\odot$ at $z \sim 1.6$ convert $\epsilon_{\rm baryon} = {\rm SFR}/\dot{M}_{\rm baryon} \sim 35\%$ of baryons accreting onto halos into new stars. This estimate is consistent with the average conversion efficiency derived by Behroozi et al. (2013a). We further give a rough estimation on the scatter in $\epsilon_{\rm baryon}$, with an assumption that the scatter in SFR is caused by independent fluctuations in $\epsilon_{\rm baryon}$ and $\dot{M}_{\rm baryon}$, which is now considered to be proportional to halo mass. Because the scatter in SFR is approximately 0.3 dex and the scatter in halo mass is expected to be 0.2 dex (see Appendix C) at fixed stellar mass ($=M_*^{\rm lim}$), the scatter in $\epsilon_{\rm baryon}$ is to be $\sim$0.2 dex (a factor of 1.6) or less.

At $M_h = M_{\rm min}$, the mass ratio $M_*/M_h$ is likely to be proportional to $M_h^{1.5}$, i.e., $M_* \propto M_h^{2.5}$ (Martinez-Manso et al. 2015; Behroozi et al. 2013a). If the baryon conversion occurs at a constant rate for all halo masses, the stellar mass grows proportionally to the halo mass, as the baryon accretion rate is expected to be almost proportional to halo mass, i.e., $\dot{M}_{\rm baryon} \propto M_h$. Therefore, the stronger dependence of $M_*$ to $M_h$ indicates that the integrated past baryon conversion depends on the halo mass. We recall that the slope of the star-forming main sequence ($M_*$–SFR) is $\sim$0.7–0.8 (e.g., Kashino et al. 2013). Given these observational facts, we find that the baryon

conversion efficiency increases moderately with increasing stellar mass, as $\epsilon_{\rm baryon} \propto M_*^{0.3-0.4}$, and more strongly with halo mass, as $\epsilon_{\rm baryon} \propto M_{\rm h}^{0.75-1.0}$. A similar dependence has also been found at a higher redshift (see Figure 13 of Harikane et al. 2016). We note that these relations pertain to the stellar mass between $10^{9.57} M_\odot$ and $M_* \sim 10^{10.5} M_\odot$, or halo mass from $M_{\rm min} = 10^{11.87} M_\odot$ to $M_{\rm h} \sim 10^{12.2} M_\odot$. Because the slope of the SHMR decreases with increasing (both stellar and halo) mass (eventually its slope becomes negative), the conversion efficiency has a peak value ($\epsilon_{\rm baryon} \sim 0.6$) that begins to decrease toward the highest masses (see Figure 11 of Behroozi et al. 2013a).

### 7.4. Satellite Galaxies

The detection of the one-halo term in the correlation function enables us to estimate the fraction of galaxies that are satellites ($f_{\rm sat}$). This fraction is dependent on $M_{\rm min}$, $M_1$, and the halo mass function. For our sample, we find a satellite fraction $f_{\rm sat} = 0.39^{+0.15}_{-0.20}$ from Equation (32). This is similar to the value of $f_{\rm sat} \approx 0.3$ for local galaxies (Zehavi et al. 2011). Guo et al. (2014) find a strong correlation between $f_{\rm sat}$ and the number density of the galaxy population, with declining $f_{\rm sat}$ toward low number density (see their Figure 5) by using a compilation of measurements at $z \lesssim 1$. The relationship they derived gives $f_{\rm sat} \sim 20\%$ for the number density of our sample ($n = 9.3 \times 10^{-3} \, h^3 \, {\rm cMpc}^{-3}$), which is broadly consistent with our result. At higher redshifts, Wake et al. (2011) measure $f_{\rm sat} \approx 0.28$ for stellar mass limit samples ($>10^{10} M_\odot$) at $z \sim 1.5$. Martinez-Manso et al. (2015) also report a similar value ($f_{\rm sat} \sim 0.25$) at $z \sim 1.5$. Our result is consistent with these values as well. Although our data do not constrain the redshift evolution of $f_{\rm sat}$ with the given statistical error, Coupon et al. (2012) suggest that the satellite fraction moderately increases with cosmic time since $z \sim 1$, which can be straightforwardly interpreted as a consequence of average growth of over-dense regions.

The ratio $M_1/M_{\rm min}$ determines the "shoulder" in $\langle N|M_{\rm h}\rangle$ (Conroy et al. 2006), or the gap between halo masses at which halos acquire a central galaxy and host an additional satellite galaxy (referred to as the "hosting gap" by Zehavi et al. 2011). Zehavi et al. (2011) found $M_1/M_{\rm min} \approx 17$ for local galaxies, and that the ratio decreases with increasing threshold luminosity of a given sample. This trend is also shown clearly by Guo et al. (2014), using a compilation of measurements, where the $M_1/M_{\rm min}$ ratio decreases with a decreasing galaxy number density. Here, we find the ratio to be $M_1/M_{\rm min} = 4.4^{+4.3}_{-2.1}$ for our FMOS sample. This is smaller than the local value, whereas our result is in good agreement with previous findings at intermediate redshifts. For example, Wake et al. (2011) find $M_1/M_{\rm min} \approx 6$ for samples with stellar mass above $\sim 10^{10} M_\odot$ at $1 \lesssim z \lesssim 2$. Recently, Martinez-Manso et al. (2015) also found similar values $M_1/M_{\rm min} \approx 7$. Conroy et al. (2006) find a decrease in $M_1/M_{\rm min}$ with increasing redshift, which may reflect the increasing fraction of relatively low-mass halos that host satellites in addition to a central galaxy. The authors argue that pairs of galaxies within the same halo with mass close to $M_{\rm min}$ predominantly contributes to the one-halo term of the correlation function at higher redshifts.

### 7.5. Prediction for the Future Subaru/PFS Survey

Finally, we give predictions for the performance of the clustering measurements expected for the future Subaru galaxy survey with the Prime Focus Spectrograph (PFS; Tamura et al. 2016). This is a next-generation multi-object spectrograph that follows FMOS, having a field of view six times larger (1.3 degree in diameter) and many more fibers ($N_{\rm fiber} = 2400$) than FMOS, as well as an unprecedented wavelength coverage (0.38–1.26 $\mu$m). The preliminary PFS survey strategy is described in Takada et al. (2014), although the design has been modified in part to date. The PFS galaxy evolution survey aims to have a dedicated program to observe roughly half a million color-selected galaxies at $1 < z < 2$, to a limiting magnitude of $J_{\rm AB} \approx 23.3$ (or deeper) over $\sim 15 \, {\rm deg}^2$. Galaxies will be selected from the Hyper Suprime-Cam Subaru Strategic Program (HSC-SSP) that provides deep and wide imaging with $grizY$ photometry.

Here, we aim to evaluate the statistical errors and how strongly the parameters can be constrained with a PFS-like sample, while assuming our best-fit values from the HOD modeling as fiducial. The signal-to-noise ratio of an observed correlation function is approximately proportional to the square root of the galaxy pair counts in bins of the separation. For an ideal survey with a contiguous survey volume and uniform sampling, the pair counts are approximately proportional to the survey volume and square of the galaxy number density. If the survey field is separated into distinct smaller subregions, the number of galaxy pairs, especially with large spatial separations, increases more slowly than expected. The PFS survey will cover multiple distinct areas, each having $\sim 6$–7 deg$^2$. Such separate regions will not significantly affect the scales that we cover in this work ($\lesssim 20$ cMpc). Therefore, we simply scale the statistical errors on the observed correlation function to match the expected survey volume and number density for the PFS survey.

For this purpose, we adopt the same redshift range as this work, i.e., $1.43 \leqslant z \leqslant 1.74$, resulting in a comoving volume of $1.75 \times 10^7 \, (h^{-1} \, {\rm cMpc})^3$ for a nominal survey area of 15 deg$^2$, which is 20 times larger than our FMOS survey. We assume a mass complete sample above $10^{10} M_\odot$ with a uniform sampling rate of 75%. Based on the COSMOS catalog (Ilbert et al. 2013), the sample size should be $N \sim 6.5 \times 10^4$, with a galaxy number density of $4 \times 10^{-3} \, (h^{-1} \, {\rm cMpc})^{-3}$, which is seven times higher than our FMOS-spec-$z$ sample (see Table 2). As a very rough estimate, the errors on the correlation function will decrease by a factor of 30 for the full sample with the same binning in $r_{\rm p}$, whereas if we divide the full sample equally into 10 subsamples, the errors become smaller by a factor of three, as compared to our data. The values of $\log w_{\rm p}(r_{\rm p})$ are then adjusted by adding statistical noise corresponding to the scaled errors. Again, we do not intend here to make any predictions for the shape and/or absolute values of correlation function, only to demonstrate the improvement in statistical accuracy.

For these two cases (the full sample with a 75% sampling rate and the 1/10 binned subsample), we show in Figure 21 the expected measurements and parameter constraints, compared to our results. For simplicity, we assume pair counting with the same ($r_{\rm p}$, $\pi$) grid, employ the same priors as in our work with the FMOS data, and use the same constraint on the galaxy number density $n_{\rm tot}$ as the real data. As mentioned in Section 6.3, when the errors on $w_{\rm p}(r_{\rm p})$ are large, a *negative* correlation appears between $M_{\rm min}$ and $M_1'$, due to the constraint
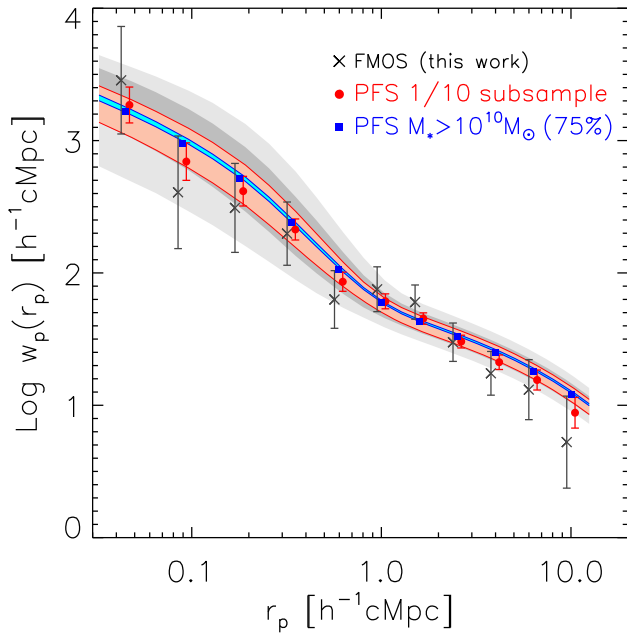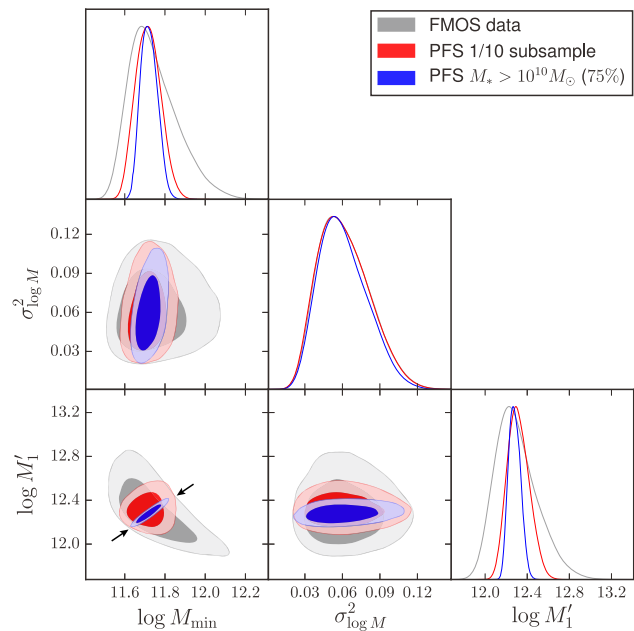
**Figure 21.** Left: correlation functions for our data and hypothetical PFS samples, as labeled. Dark and light gray regions show the 68 and 95% confidence intervals of the HOD model for our data; light red and cyan regions indicate the 95th percentiles for the PFS samples. Right: the constraints of the HOD parameters for each sample.

on $n_{\rm tot}$ (see gray contours); the widths of the contours (between two arrows in the right panel of Figure 21) are determined by the uncertainty on the abundance constraint. In contrast, the shape and amplitude of $w_{\rm p}(r_{\rm p})$ control $M_{\rm min}$ and $M_1'$, while allowing *positive* correlation between them. Hence, when the errors on $w_{\rm p}(r_{\rm p})$ are very small, the positive correlation appears in a range permitted by the constraint on $n_{\rm tot}$, as shown by blue contours in Figure 21.

As is evident, the statistical uncertainties are remarkably improved for the hypothetical samples for a future survey, even if the samples are limited to a narrow range of redshift. For the full sample, given the levels of their statistical errors, it is obvious that the systematic effects that we have discussed in this paper, i.e., the impact of fiber allocation and/or inhomogeneous line-of-sight detection rate, will significantly influence the measurements—and thus the physical interpretations. In the right panel of Figure 21, it is shown that the constraints on $M_{\rm min}$ and $M_1'$ are significantly improved, whereas those on $\sigma_{\log M}^2$ are almost identical to our FMOS result. In fact, this parameter is essentially determined by the prior. This means that even very statistically accurate clustering measurements cannot strongly constrain all parameters. Therefore, alternative methods are still necessary to fully understand the galaxy-halo relation. In conclusion, the future multi-object spectroscopic surveys (e.g., PFS, MOONS, DESI) assure significant improvement in the statistics with respect to clustering measurements, whereas systematic effects must be excluded more carefully. Precise models are imperative to interpret such upcoming observational data.

## 8. Summary

We have investigated the clustering properties of star-forming galaxies at $1.43 \leqslant z \leqslant 1.74$, using the data set from the FMOS-COSMOS survey. With 516 galaxies having an H$\alpha$ detection, we measured the projected two-point correlation

function, and investigated the properties of dark matter halos through the HOD modeling of galaxy clustering. Our main results are as follows.

1. The observed correlation function indicates a significant clustering at $0.04 \lesssim r_{\rm p}/(h^{-1}\,{\rm cMpc}) \lesssim 10$. By modeling with a power-law function, we find a correlation length of $r_0 = 5.26^{+0.75}_{-0.62}\,h^{-1}\,{\rm cMpc}$, which is consistent with preceding studies using galaxy samples with stellar masses similar to our study.

2. We model the observed correlation function using an HOD model (Equations (29) and (30)). The HOD parameters are effectively constrained with our current data, and a significant one-halo term is confirmed with a transition scale of $r_{\rm p} \simeq 0.5\,h^{-1}\,{\rm cMpc}$, where the one- and two-halo terms are equivalent.

3. We derive an effective large-scale bias, $b_{\rm eff} = 2.18^{+0.18}_{-0.24}$, and find that star-forming galaxies with $M_* \geqslant 10^{9.57}\,M_\odot$ reside, on average, in halos with mass $M_{\rm eff} = 4.71^{+1.19}_{-1.62} \times 10^{12}\,h^{-1}\,M_\odot$, which is consistent with other measurements from HOD modeling. With predictions of the mass assembly histories, we find that these halos will have grown into group-scale halos ($\sim 2 \times 10^{13}\,h^{-1}\,M_\odot$) at the present epoch.

4. Our work investigates the low-mass part of the stellar-to-halo mass relation at this redshift, with a new constraint, $\log M_*/M_{\rm h} = -2.3 \pm 0.2$ at $M_{\rm h} = 10^{11.87}\,M_\odot$, while confirming the decline in the stellar-to-halo mass ratio at $M_{\rm h} < 10^{12}\,M_\odot$. However, there are discrepancies between SHMRs based on different observations and/or analyses. In particular, SHMRs based on HOD modeling, including our result, seem to have systematically lower values at lower halo masses than those based on abundance matching.

5. We find the efficiency of converting baryons into stars relative to the total amount of baryons accreting onto

halos to be ~35%, with a maximum scatter of a factor of ~1.6. The baryon conversion efficiency $\epsilon_{baryon}$ depends on stellar and halo mass, as $\epsilon \propto M_*^{0.3-0.4}$ and $\epsilon \propto M_h^{0.75-1.0}$, respectively, around the threshold mass up to the peak mass of the SHMR ($\log M_*/M_\odot \approx$ 9.57–10.5, $\log M_h/M_\odot \approx 11.87$–12.2).

6. We find the satellite fraction to be $f_{sat} = 0.39^{+0.15}_{-0.20}$, which is consistent with other measurements at both local and higher redshifts. In addition, the $M_1/M_{min}$ ratio is found to be ~4.4, which is lower than typically seen for low-$z$ galaxies ($\simeq 17$ at $z \lesssim 0.1$), but consistent with other studies at higher redshifts ($z > 1$), which show $M_1/M_{min} < 10$. These results suggest that galaxy pairs in the same, relatively low-mass halos significantly contribute to the one-halo term of the correlation function.

In addition, we established correction schemes for fiber allocation and inhomogeneous detection due to sky contamination, and demonstrated the effectiveness of these methods even for small scales (i.e., the one-halo regime). Such techniques will be effective in measuring the intrinsic galaxy clustering in future spectroscopic surveys with next-generation multi-object spectrographs, such as Subaru/PFS. Future surveys will provide a sample of ~$5 \times 10^5$ galaxies at $z > 1$, and allow us to measure galaxy clustering as a function of galactic properties at much higher statistical accuracy. Therefore, the systematic effects must be excluded carefully, and precise models are required to maximally draw information from such upcoming observational data.

# Appendix A
## Correction for Critical Biases in the Observed Correlation Function

### A.1. Mock Samples from the $\nu^2$ GC Simulation

We construct mock samples to evaluate our correction scheme for the effects of fiber allocation and sky contamination, using the new numerical Galaxy Catalog ($\nu^2$GC) cosmological simulation (Ishiyama et al. 2015). We use the mid-size simulation ($\nu^2$GC-M) that was conducted using $4096^3$ particles with a mass of $2.20 \times 10^8 \, h^{-1} \, M_\odot$, in a comoving cube with a side length of $560 \, h^{-1}$ cMpc. We employ a halo/subhalo catalog at a scale factor $a = 0.384871$ ($z = 1.598$), which is close to the median redshift of our spectroscopic sample. Halos and subhalos are identified using the Rockstar algorithm (Behroozi et al. 2013b), resolving subhalos with $M_h \sim 10^{10} \, h^{-1} \, M_\odot$. The lengths of our survey volume in the

plane of the sky and along the line of sight are ~$60 \, h^{-1}$ cMpc and ~$400 \, h^{-1}$ cMpc over $1.43 \leqslant z \leqslant 1.74$, respectively. Hence, we can divide the simulation box into 64 sub-boxes on the $x$-$y$ plane, each having a volume of $70 \times 70 \times 560 \, (h^{-1} \, \text{cMpc})^3$, that can enclose the survey volume. In doing so, we implicitly ignore the redshift evolution between this redshift range. We then convert the comoving coordinates ($x$, $y$, $z$) into angular coordinates and redshift, associating the middle point in a box to the center of the survey volume. We take the peculiar velocity along the $z$-axis of individual (sub) halos into account, to derive the true redshift (as opposed to the cosmological redshift).

For each sub-box, we select halos and subhalos from the catalog to create a sample that has a number density similar to the FMOS-parent sample ($n = 3.07 \times 10^{-3} \, h^3 \, \text{cMpc}^{-3}$) and a correlation function whose amplitude is also similar to the observed one (see Figure 24). We extract halos that have a maximum circular velocity $V_{max}$ greater than $170 \, \text{km s}^{-1}$ and subhalos that have $V_{max} > 160 \, \text{km s}^{-1}$. It is known that the highest $V_{max}$ that a halo can attain over its past history (often called $V_{peak}$) is a better proxy of stellar mass, rather than $V_{max}$ at the epoch of interest, especially in order to better reproduce galaxy clustering (Reddick et al. 2013). This is reflective of the fact that subhalos lose their mass and $V_{max}$ when they accrete onto their parent halos. However, such information is not yet available for this simulation. To roughly account for the mass loss of subhalos, we here use a lower threshold of $V_{max}$ for subhalos than for parent halos. However, the precise selection does not matter because we are only interested in the relative change due to observational biases, but not in any absolute quantities predicted by the simulation.

To ensure that the redshift distribution of the mock samples matches that of the FMOS-parent sample, we assign *photometric* redshifts with additional random fluctuations of rms ($\Delta z$) = 0.062 to the true redshift of each halo. We then extract objects from the same range of the photometric redshift as we applied for the real data ($1.46 \leqslant z_{phot} \leqslant 1.72$). From them, we randomly select ~2300 objects to match the number of the FMOS-parent sample, and refer to these as the *Mock-parent* samples. In addition, a subset of the objects having a true redshift $1.43 \leqslant z_{true} \leqslant 1.74$ is subtracted to make the *Mock-par-specz* samples, whose average correlation function is used as an intrinsic quantity to be compared with biased measurements. From these samples, we construct two types of mock samples that take into account the inhomogeneous detection along the line-of-sight (mainly due to the atmospheric lines) and fiber allocation effects.

We randomly select the same number of objects as present in our FMOS-spec-$z$ sample (516) from each of the Mock-par-spec$z$ samples. During this selection, we account for the weight as a function of redshift, as shown in Figure 7, to reflect the non-uniform detectability of the H$\alpha$ emission line along the line of sight. These *Mock-zweight* samples are used in the subsequent section.

We next prepare mocks to assess the impact of fiber allocation algorithm. For each Mock-parent sample, we mimic the selection function induced by the FMOS fiber allocation software to construct the *Mock-fiber-target* samples. We also account for fibers that broke down; some of these are regularly out of order, and others depend on the observing runs. These samples have approximately the same number of galaxies (~1200) as the FMOS-fiber-target sample, representing

**Table 5**
Mock Samples from the $\nu^2$GC Simulation

| Sample name | N | Conditions[a] |
|---|---|---|
| Mock-parent | ∼2300 | $1.46 \leqslant z_{\rm phot} \leqslant 1.72$[b] |
| Mock-par-specz | ∼2100 | $1.46 \leqslant z_{\rm phot} \leqslant 1.72$, and $1.43 \leqslant z_{\rm true} \leqslant 1.74$[c] |
| Mock-zweight | 516 | $1.46 \leqslant z_{\rm phot} \leqslant 1.72$, $1.43 \leqslant z_{\rm true} \leqslant 1.74$, and non-uniform detectability[d] |
| Mock-fiber-target | ∼1200 | $1.46 \leqslant z_{\rm phot} \leqslant 1.72$ and fiber allocation |
| Mock-fiber-specz | 516 | $1.46 \leqslant z_{\rm phot} \leqslant 1.72$, fiber allocation, and $1.43 \leqslant z_{\rm true} \leqslant 1.74$ |

**Notes.**
[a] The thresholds on $V_{\rm max}$ and the realistic shape of the FMOS survey area are applied for all the mock samples listed here.
[b] Photometric redshift of a (sub)halo is generated by addition of a random error to the true redshift.
[c] Spectroscopic redshift $z_{\rm true}$ of a (sub)halo includes the effect of peculiar motion.
[d] The inhomogeneous weight function shown in Figure 7 is applied.

the observed (fiber-allocated) galaxies. We then create the *Mock-fiber-specz* samples by randomly extracting 516 objects with a true redshift between $1.43 \leqslant z \leqslant 1.74$ from each Mock-fiber-target catalog. These final samples correspond to the FMOS-spec-$z$ sample that is used for clustering measurement. We note that the fiber allocation is not considered for the Mock-$z$weight samples, and the inhomogeneous detection rate is not taken into account for the Mock-fiber-target/spec-$z$ samples to examine each of these effects individually. The selections of these mock samples are summarized in Table 5.

### A.2. Inhomogeneous Line-of-sight Detectability

Inhomogeneous detection of the H$\alpha$ emission line along the redshift direction, due to the OH lines/mask and instrumental characteristics can induce an artificial clustering signal along that direction, parallel to the line of sight. As a result, the amplitude of the correlation function is expected to be artificially enhanced if corrections are not applied. In our analysis, we used a random sample wherein redshifts were sampled from the intrinsic distribution of the FMOS-parent sample, but with a probability weighted by the non-uniform detectability as a function of redshift (Figure 7).

Let us now demonstrate the effects and reliability of our method. We calculate the correlation functions of a series of Mock-$z$weight samples, in which the non-uniform detectability is taken into account by selecting 516 objects with the weighted probability (Appendix A.1). To assess the effect on our calculation of the correlation function, using Equation (2), we inspect differences in the clustering by using random samples with and without the non-uniform detection rate corrections. In Figure 22, the average correlation function of 64 mock samples is shown, corresponding to the use of these random catalogs, compared to the intrinsic correlation function of the sample. It can be seen that the use of the modified random samples accurately recovers the intrinsic one (filled red circles). In contrast, the values of $w_{\rm p}(r_{\rm p})$ are slightly—but systematically—enhanced over the range of $r_{\rm p} > 0.1\,h^{-1}$ cMpc, if we do not consider the non-uniform weight in the random sample (open blue circles). This enhancement of the clustering amplitude results in an overestimate of the correlation length by $\Delta r_0 \sim 0.2$–$0.3\,h^{-1}$ cMpc. Although this systematic error is well below the uncertainty on the measurement from our current data, it could be critical in future studies based on large samples, which provide much higher statistical accuracy than our study (see Section 7.5). This exercise evidently indicates that the effects of non-uniform detectability along the redshift
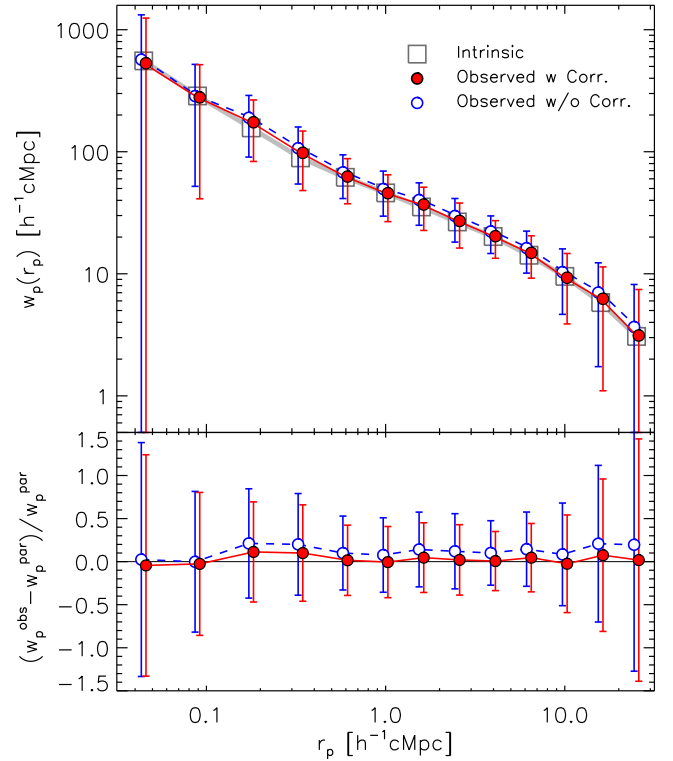


**Figure 22.** Upper panel: projected two-point correlation functions from the mock samples. The average *observed* $w_{\rm p}(r_{\rm p})$ is calculated for the Mock-$z$weight samples with (filled red circles) and without (blue empty circles) the correction in the random sample for the non-uniform detection rate along the line of sight. The $x$-axis values of these data points are slightly shifted, for the purpose of display. Large squares indicate the intrinsic correlation function computed for the Mock-parent-$z_{\rm spec}$ samples. Lower panel: difference between the intrinsic and observed $w_{\rm p}(r_{\rm p})$ with or without the correction. Symbols are the same as in the upper panel.

direction can be mitigated by applying an appropriate weight function for the random sample.

### A.3. Fiber Allocation

The selection of targets is dependent on the software that allocates fibers to objects (Akiyama et al. 2008). As a result, there can be artificial biases in the on-sky distribution of objects, because not all galaxies in the input catalog are observed. After allocating a pair of fibers to one galaxy, the
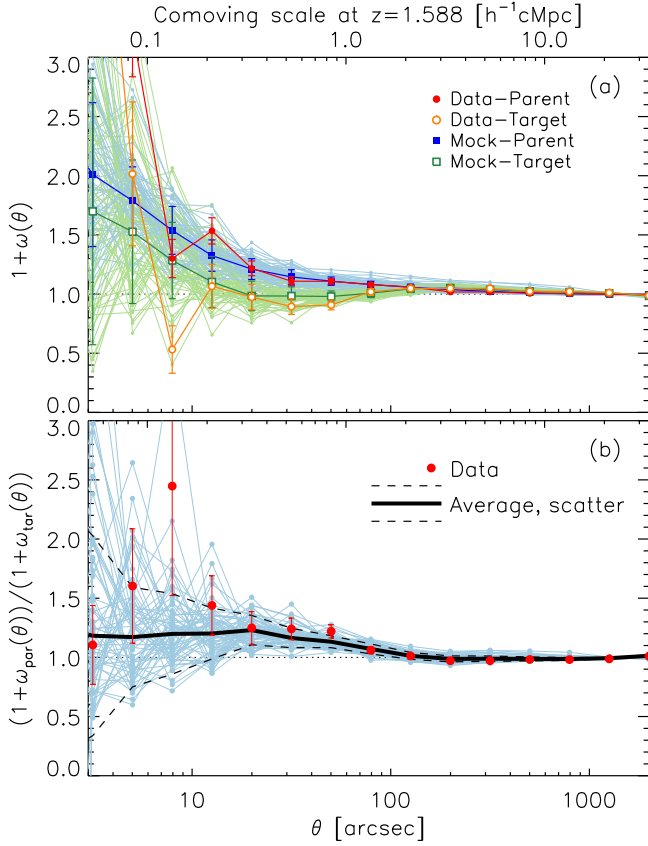
**Figure 23.** Panel (a): angular correlation function $\omega(\theta)$ for the data and the mock samples. The corresponding comoving separation at $z = 1.588$ is shown in the upper x-axis. Filled red and open orange circles show $1 + \omega(\theta)$ for the FMOS-parent and FMOS-fiber-target samples, respectively. Filled blue and open green squares show the average values of the 64 Mock-parent and Mock-fiber-target samples, respectively, with the individual measurements (light color dots). There is a suppression at scales of $\theta \lesssim 100$ arcsec due to the fiber allocation. Panel (b): ratios of $1 + \omega(\theta)$ between the parent and fiber-target samples for the data (red circles) and the individual mocks (light blue dots). The average of the mocks, which is used as the weight function for the analysis, and the scatter for the mock samples are indicated by thick solid and thin dashed lines, respectively.



**Figure 24.** Symbols are the same as in Figure 22, but the average *observed* $w_p(r_p)$ are calculated for the Mock-fiber-specz samples with and without the correction for the effects of fiber allocation.

opportunity for its neighboring galaxies to be observed at the same time decreases, due to the lack of fibers and to avoid fiber entanglement. In contrast, galaxy pairs with separations of several times the fiber separation are easy to simultaneously assign as targets. This results in a bias in the pair counts—hence the observed correlation function. In addition, the sampling rate varies across the survey area, because of the different number of exposures among the four footprints and the existence of the small areas where the footprints overlap. Moreover, some broken fibers result in holes where the sampling rate is reduced. In order to correct for these biases, we apply a weight for each galaxy pair as a function of their angular separation (see, e.g., de la Torre et al. 2011; Durkalec et al. 2015b). The weight is defined by Equation (12), using the ratio of the angular correlation function of the input catalog and that of the sample of galaxies for which fibers were allocated.

In Figure 23, we demonstrate the effects of fiber allocation. The angular correlation functions $\omega(\theta)$ of the real data, FMOS-parent, and FMOS-fiber-target samples are shown in panel (a). It is evident that the angular clustering amplitude of the FMOS-fiber-target sample is suppressed, compared to the FMOS-
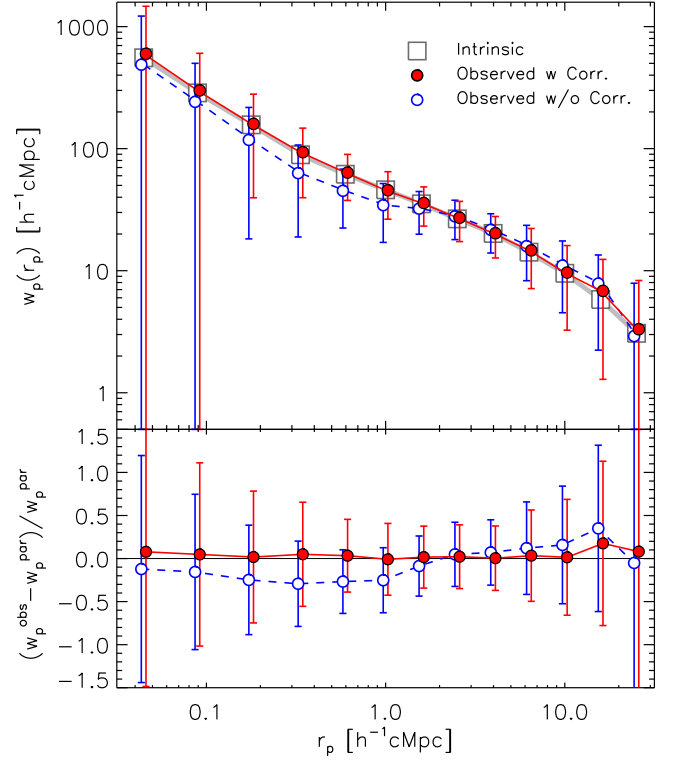
parent sample at scales $\theta \lesssim 100$ arcsec. This angular scale is similar to the fiber separation, corresponding to $r_p \lesssim 2\,h^{-1}$ cMpc at $z \sim 1.6$. We also show the angular correlation functions for 64 sets of Mock-parent and Mock-fiber-target samples. Although the scatter is large, the suppression at small scales is evident in their average values, in agreement with the real data. This decrease in the angular clustering amplitude is expected to result from the fiber allocation. Figure 23(b) shows the ratios of $1 + \omega(\theta)$ between the parent and fiber-target samples for both the data (red circles) and mocks (light blue dots). For our analysis, we use the average ratios calculated from the mock samples (thick solid line) as the weight function (Equation (12)). The weight function is >1 at angular scales of $\theta \lesssim 100$ arcsec and reaches $\sim 1.2$ at $\theta < 20$ arcsec; it is slightly smaller than unity at scales of several hundreds of arcsec. On a large scale, it reflects the ease of observing pairs with separation several times larger than the fiber separation.

Now we examine the effectiveness of our correction scheme. In Figure 24, we compare the average correlation functions $w_p(r_p)$ of 64 Mock-fiber-specz samples, with and without correction to the intrinsic correlation function. The average amplitude of the Mock-fiber-specz samples without correction is reduced at scales $r_p \lesssim 2\,h^{-1}$ cMpc, whereas it is slightly enhanced at $r_p \gtrsim 3\,h^{-1}$ cMpc, as clearly seen in the lower panel. In contrast, application of the correction scheme recovers the true $w_p(r_p)$ over the entire scale of interest, in a relatively unbiased manner. These biases are not negligible, even for studies using a relatively small sample (such as ours) and necessarily will have a more significant impact for studies with larger samples. We conclude that the proposed correction scheme based on the angular correlation functions works

appropriately in our analysis (and should work for future multi-object surveys) even for small scales down to $r_{\rm p} \sim 0.1 \, h^{-1}$ cMpc.

## Appendix B
## The Effects of Sample Selection

Here, we assess the effect of stellar mass completeness on the constraints of our HOD model. For this purpose, we construct mock samples using the Bolshoi simulation (Klypin et al. 2011), to take the advantage of its better mass resolution and the availability of merger histories of halos, instead of the $\nu^2$GC simulation used above. The simulation consists of $2048^3$ particles with a mass of $1.35 \times 10^8 \, h^{-1} \, M_\odot$ in a comoving cube with a side of $250 \, h^{-1}$ cMpc. The volume of the simulation box is 19 times larger than our survey volume, but the side length is shorter than the radial extend of the survey ($\sim 380 \, h^{-1}$ cMpc). We utilize the public catalog at $a = 0.38435$ ($z = 1.602$), in which halos are identified down to a mass of $10^{10} \, h^{-1} \, M_\odot$ based on the Rockstar halo-finding algorithm (Behroozi et al. 2013b). For the following exercise, we use the same cosmology as the Bolshoi simulation ($\Omega_m, \Omega_\Lambda$) = (0.27, 0.73).

For abundance matching, we use the peak maximum circular velocity ($V_{\rm peak}$) provided in the public catalog as an indicator of the stellar mass of a galaxy hosted by a (sub)halo. Assuming a tight correlation between $V_{\rm peak}$ and stellar mass, with no scatter, we chose the lower limit $V_{\rm peak}^{\rm lim}$ by matching the (sub)halo number density $n(>V_{\rm peak})$ to the number density of the $M_*$-selected galaxy sample (Table 2). Note that the number density is recomputed, as $n = 8.47 \times 10^{-3} \, (h^{-1} \, {\rm cMpc})^{-3}$, based on the above cosmology. In total, 132,408 (sub)halos with $V_{\rm peak} > V_{\rm peak}^{\rm lim} \approx 165 \, {\rm km \, s^{-1}}$ are selected. For each (sub) halo, a stellar mass is assigned, while keeping the relation $n(>V_{\rm peak}) = n(>M_*)$, in order to reproduce the stellar mass function of the $M_*$-selected sample (see Figure 3). We denote this mock catalog as *Mock-$V_{\rm peak}$-limited* sample. From this sample, we construct the *Mock-$M_*$-incomplete* sample, in which the stellar mass incompleteness is taken into account, by randomly selecting (sub)halos with a probability as a function of the assigned stellar masses, such that the resulting stellar mass function equals that of the FMOS-parent sample (Figure 3). In total, 43,777 (sub)halos are selected.

We measure the projected correlation function of these selected halos. Figure 25 compares the correlation functions for the Mock-$V_{\rm peak}$-limited (red triangles), Mock-$M_*$-incomplete (blue squares), and FMOS-spec-$z$ (black circles) samples. We correct the correlation functions of the mock samples for the effect of the finite size of the simulation box (see van den Bosch et al. 2013 for details). We also note that the correlation function for the real data is multiplied by $1/(1 - f_{\rm fake})^2$, where $f_{\rm fake} = 0.099$, which is the expected value from the HOD fitting (Table 4). It can be seen that the clustering amplitude of the Mock-$M_*$-incomplete sample (blue squares) is slightly enhanced relative to the Mock-$V_{\rm peak}$-limited sample (red triangles). This is naturally expected from the fact that the $M_*$-incomplete sample is, on average, biased toward massive halos, which are more strongly clustered. We fit the HOD model (Equations (29) and (30)) to the correlation functions of these two mock samples. Here, we scale the covariance matrix of the mock measurements by a factor of 19, which is the ratio of the simulation box and survey, to match the level of the statistics to the real data. The parameter constraints are
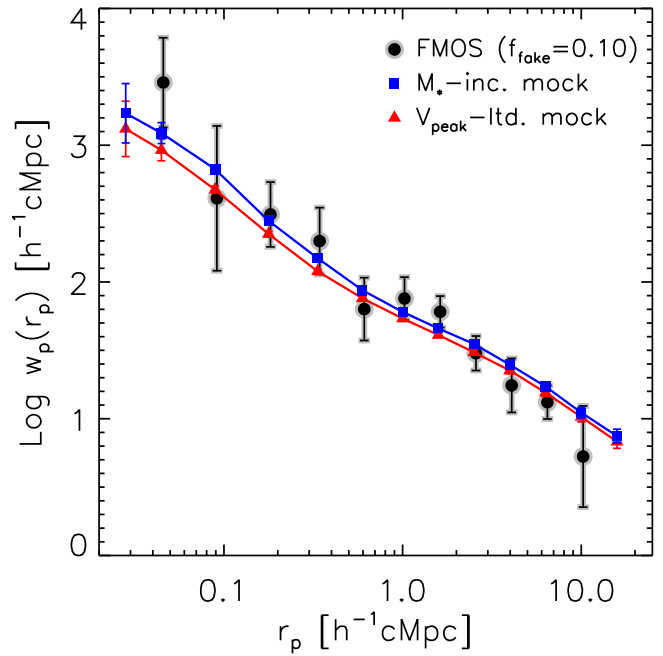


**Figure 25.** Projected two-point correlation functions for the FMOS-spec-$z$ (circles), Mock-$V_{\rm peak}$-limited (red triangles), and Mock-$M_*$-incomplete samples (blue squares).

shown separately in Figure 26 for the $V_{\rm peak}$-limited and $M_*$-incomplete samples, compared to the results for the real FMOS sample. We find that all parameters are consistent within their $1\sigma$ confidence level, although the peak positions of the posterior distributions are slightly different among data and mocks. In particular, the constraints are almost identical for the two mock samples. Theses results indicate that the stellar mass incompleteness in our FMOS sample should not have a significant impact on our conclusions, given the current level of statistical errors.

## Appendix C
## Limitation on $\sigma_{\log M}$

In this section, we present the prior that we assign to the HOD parameter, $\sigma_{\log M}$ (see Section 6.2). We use the SHMR derived by Behroozi et al. (2013a), who parameterized its evolution and intrinsic scatter over $0 < z < 8$, and constrained it using a variety of observational constraints. Although our constraint on the stellar-to-halo mass ratio deviates from the Behroozi et al. (2013a) SHMR (see Section 7.2), this discrepancy is not significant in the following calculations, because only the scatter and the slope of the SHMR matter for the prior that we use for $\sigma_{\log M}$.

The scatter in stellar masses ($\log M_*$) of galaxies that inhabit halos of a given mass is constrained to be $\sigma_{\log M_*}^{\rm int} = 0.23 \pm 0.04$ at $z = 1.588$. In addition, we incorporate another scatter ($\sigma_{\log M_*}^{\rm sample}$) that accounts for sample selection. We select galaxies based on the *best* estimate of the stellar mass from the SED, which should differ from the *true* value due the uncertainties in the photometric redshift, flux measurement, and population synthesis model. Therefore, a threshold in the estimated stellar mass does not correspond to a threshold in the true stellar mass, and one can expect low stellar mass objects to scatter into our sample. The width of this scatter will
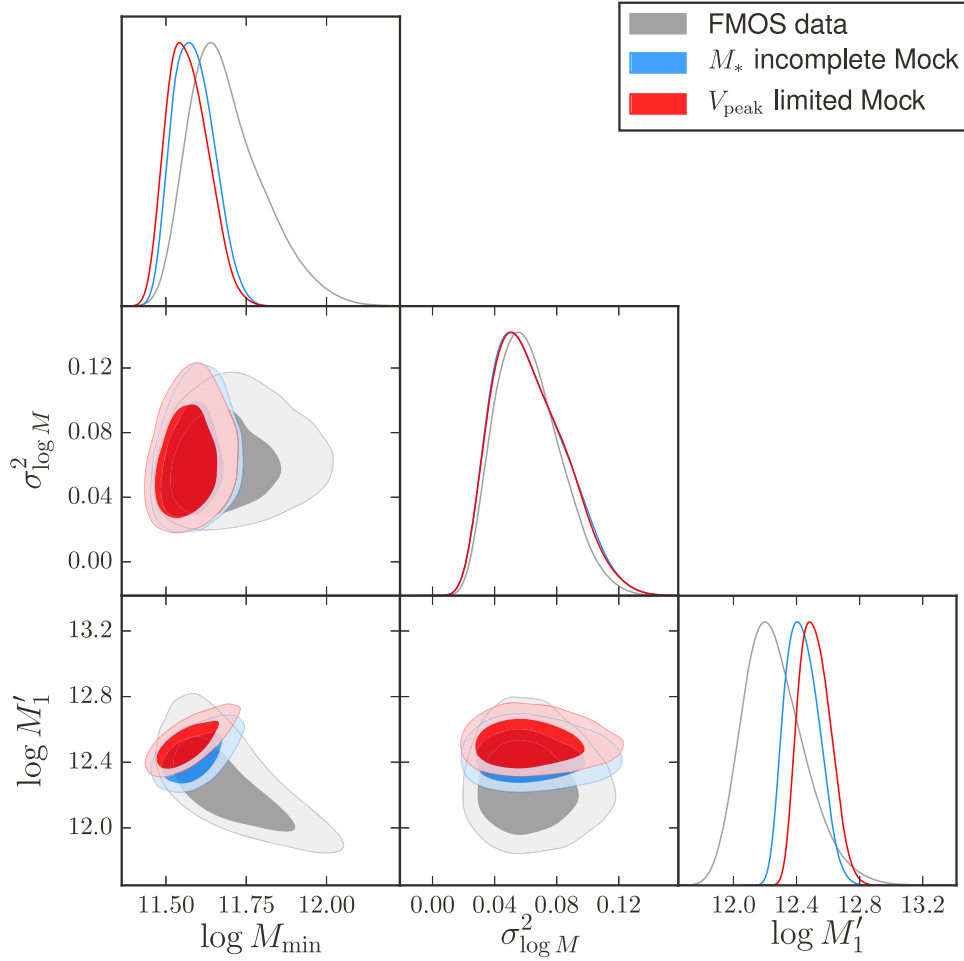
**Figure 26.** Constraints of the HOD parameters ($M_{\min}$, $\sigma^2_{\log M}$, $M'_1$) for the Mock-$V_{\mathrm{peak}}$-limited (red) and Mock-$M_*$-incomplete samples (blue), compared to the FMOS results (gray).

correspond to the typical error in the stellar mass estimation. We compute for each galaxy the probability that its true stellar mass satisfies $M_* \geqslant M^{\mathrm{lim}}_*$, and derive the probability-weighted stellar mass distribution of our sample such that it has a smoothed lower limit at $M_* = M^{\mathrm{lim}}_* \equiv 10^{9.57}\,M_\odot$. We find $\sigma^{\mathrm{sample}}_{\log M_*} = 0.09$ by deconvolving the distribution into a sharp-cutoff function and a Gaussian function with a standard deviation $\sigma^{\mathrm{sample}}_{\log M_*}$. The total scatter in $\log M_*$ of galaxies within halos of the threshold mass is thus given as

$$\sigma_{\log M_*} = \sqrt{(\sigma^{\mathrm{int}}_{\log M_*})^2 + (\sigma^{\mathrm{sample}}_{\log M_*})^2} = 0.24 \pm 0.04. \quad (36)$$

We then calculate the probability of finding a galaxy of $M_* = M^{\mathrm{lim}}_*$ in a halo as a function of halo mass:

$$P(\log M^{\mathrm{lim}}_* | M_{\mathrm{h}}) = \frac{1}{\sqrt{2}\,\pi\sigma_{\log M_*}}$$
$$\times \exp\left(-\frac{(\log M^{\mathrm{lim}}_* - \log M_*(M_{\mathrm{h}}))^2}{2\sigma^2_{\log M_*}}\right), \quad (37)$$

where $M_*(M_{\mathrm{h}})$ is the SHMR parameterized by Behroozi et al. (2013a). The expected mean and mean square of halo mass to

have a galaxy of $M_* = M^{\mathrm{lim}}_*$ are given by

$$\langle (\log M)^m \rangle$$
$$= \frac{\int P(\log M^{\mathrm{lim}}_* | M_{\mathrm{h}})(\log M_{\mathrm{h}})^m\, n(M_{\mathrm{h}})\,dM_{\mathrm{h}}}{\int P(\log M^{\mathrm{lim}}_* | M_{\mathrm{h}})\, n(M_{\mathrm{h}})\,dM_{\mathrm{h}}}. \quad (38)$$

where $n(M_{\mathrm{h}})$ is a halo mass function at a given redshift, and $m = 1(2)$ corresponds to the mean (mean square). We then find the deviation in halo mass at our threshold stellar mass:

$$\sqrt{\mathrm{Var}(\log M_{\mathrm{h}})} = \sqrt{\langle (\log M_{\mathrm{h}})^2 \rangle - \langle \log M_{\mathrm{h}} \rangle^2}$$
$$\approx 0.17 \pm 0.02. \quad (39)$$

The HOD parameter $\sigma_{\log M}$ is then given by $\sigma_{\log M} = \sqrt{2} \times \sqrt{\mathrm{Var}(\log M_{\mathrm{h}})} = 0.24 \pm 0.3$. This is the prior information we adopt for the parameter in our HOD modeling. We have checked that the inclusion of the redshift evolution of the Behroozi et al. (2013a) SHMR over $1.43 \leqslant z \leqslant 1.74$ does not change these values. The derived value of $\sigma_{\log M}$ is consistent with an estimate in previous studies (e.g., More et al. 2011), albeit at local redshifts.

# References

Abbas, U., de la Torre, S., Le Fèvre, O., et al. 2010, MNRAS, 406, 1306
Adelberger, K. L., Steidel, C. C., Pettini, M., et al. 2005, ApJ, 619, 697
Adelberger, K. L., Steidel, C. C., Shapley, A. E., et al. 2004, ApJ, 607, 226
Akiyama, M., Smedley, S., Gillingham, P., et al. 2008, Proc. SPIE, 7018, 70182V
Alam, S., Ata, M., Bailey, S., et al. 2016, MNRAS, submitted, (arXiv:1607.03155)
Arnouts, S., & Ilbert, O. 2011, LePHARE: Photometric Analysis for Redshift Estimate, Astrophysics Source Code Library, ascl:1108.009
Behroozi, P. S., Conroy, C., & Wechsler, R. H. 2010, ApJ, 717, 379
Behroozi, P. S., Wechsler, R. H., & Conroy, C. 2013a, ApJ, 770, 57
Behroozi, P. S., Wechsler, R. H., & Wu, H.-Y. 2013b, ApJ, 762, 109
Berlind, A. A., & Weinberg, D. H. 2002, ApJ, 575, 587
Berlind, A. A., Weinberg, D. H., Benson, A. J., et al. 2003, ApJ, 593, 1
Béthermin, M., Kilbinger, M., Daddi, E., et al. 2014, A&A, 567, A103
Bielby, R., Hill, M. D., Shanks, T., et al. 2013, MNRAS, 430, 425
Bielby, R. M., Gonzalez-Perez, V., McCracken, H. J., et al. 2014, A&A, 568, A24
Brown, M. J. I., Dey, A., Jannuzi, B. T., et al. 2003, ApJ, 597, 225
Brown, M. J. I., Zheng, Z., White, M., et al. 2008, ApJ, 682, 937
Bruzual, G., & Charlot, S. 2003, MNRAS, 344, 1000
Cacciato, M., van den Bosch, F. C., More, S., et al. 2009, MNRAS, 394, 929
Cacciato, M., van den Bosch, F. C., More, S., Mo, H., & Yang, X. 2013, MNRAS, 430, 767
Calzetti, D., Armus, L., Bohlin, R. C., et al. 2000, ApJ, 533, 682
Chabrier, G. 2003, PASP, 115, 763
Coil, A. L., Davis, M., Madgwick, D. S., et al. 2004, ApJ, 609, 525
Cole, S., & Kaiser, N. 1989, MNRAS, 237, 1127
Colless, M., Dalton, G., Maddox, S., et al. 2001, MNRAS, 328, 1039
Conroy, C., Wechsler, R. H., & Kravtsov, A. V. 2006, ApJ, 647, 201
Coupon, J., Arnouts, S., van Waerbeke, L., et al. 2015, MNRAS, 449, 1352
Coupon, J., Kilbinger, M., McCracken, H. J., et al. 2012, A&A, 542, A5
Daddi, E., Cimatti, A., Renzini, A., et al. 2004, ApJ, 617, 746
Daddi, E., Dannerbauer, H., Stern, D., et al. 2009, ApJ, 694, 1517
Daddi, E., Dickinson, M., Morrison, G., et al. 2007, ApJ, 670, 156
de la Torre, S., Le Fèvre, O., Porciani, C., et al. 2011, MNRAS, 412, 825
Durkalec, A., Le Fèvre, O., de la Torre, S., et al. 2015a, A&A, 576, L7
Durkalec, A., Le Fèvre, O., Pollo, A., et al. 2015b, A&A, 583, A128
Fisher, K. B., Davis, M., Strauss, M. A., Yahil, A., & Huchra, J. P. 1994, MNRAS, 267, 927
Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013, PASP, 125, 306
Foucaud, S., McCracken, H. J., Le Fèvre, O., et al. 2003, A&A, 409, 835
Frieman, J. A., Turner, M. S., & Huterer, D. 2008, ARA&A, 46, 385
Geach, J. E., Sobral, D., Hickox, R. C., et al. 2012, MNRAS, 426, 679
Gobat, R., Daddi, E., Béthermin, M., et al. 2015, A&A, 581, A56
Guo, H., Zheng, Z., Zehavi, I., et al. 2014, MNRAS, 441, 2398
Guo, Q., White, S., Li, C., & Boylan-Kolchin, M. 2010, MNRAS, 404, 1111
Guzzo, L., Bartlett, J. G., Cappi, A., et al. 2000, A&A, 355, 1
Han, J., Eke, V. R., Frenk, C. S., et al. 2015, MNRAS, 446, 1356
Harikane, Y., Ouchi, M., Ono, Y., et al. 2016, ApJ, 821, 123
Hartley, W. G., Almaini, O., Cirasuolo, M., et al. 2010, MNRAS, 407, 1212
Hartley, W. G., Lane, K. P., Almaini, O., et al. 2008, MNRAS, 391, 1301
Hawkins, E., Maddox, S., Cole, S., et al. 2003, MNRAS, 346, 78
Hayashi, M., Shimasaku, K., Motohara, K., et al. 2007, ApJ, 660, 72
Hoekstra, H., Hsieh, B. C., Yee, H. K. C., Lin, H., & Gladders, M. D. 2005, ApJ, 635, 73
Ilbert, O., McCracken, H. J., Le Fèvre, O., et al. 2013, A&A, 556, A55
Ishikawa, S., Kashikawa, N., Toshikawa, J., & Onoue, M. 2015, MNRAS, 454, 205
Ishiyama, T., Enoki, M., Kobayashi, M. A. R., et al. 2015, PASJ, 67, 61
Iwamuro, F., Moritani, Y., Yabe, K., et al. 2012, PASJ, 64, 59
Kaiser, N. 1987, MNRAS, 227, 1
Kartaltepe, J. S., Sanders, D. B., Silverman, J. D., et al. 2015, ApJL, 806, L35
Kashikawa, N., Yoshida, M., Shimasaku, K., et al. 2006, ApJ, 637, 631
Kashino, D., Silverman, J. D., Rodighiero, G., et al. 2013, ApJL, 777, L8
Kashino, D., Silverman, J. D., Sanders, D., et al. 2017, ApJ, 835, 88
Kauffmann, G., Colberg, J. M., Diaferio, A., & White, S. D. M. 1999, MNRAS, 307, 529
Kennicutt, R. C., Jr. 1998, ARA&A, 36, 189
Kimura, M., Maihara, T., Iwamuro, F., et al. 2010, PASJ, 62, 1135
Klypin, A. A., Trujillo-Gomez, S., & Primack, J. 2011, ApJ, 740, 102
Kravtsov, A. V., Berlind, A. A., Wechsler, R. H., et al. 2004, ApJ, 609, 35
Kravtsov, A. V., Vikhlinin, A., & Nagai, D. 2006, ApJ, 650, 128

Laigle, C., McCracken, H. J., Ilbert, O., et al. 2016, ApJS, 224, 24
Landy, S. D., & Szalay, A. S. 1993, ApJ, 412, 64
Le Fèvre, O., Vettolani, G., Garilli, B., et al. 2005, A&A, 439, 845
Leauthaud, A., Tinker, J., Behroozi, P. S., Busha, M. T., & Wechsler, R. H. 2011, ApJ, 738, 45
Leauthaud, A., Tinker, J., Bundy, K., et al. 2012, ApJ, 744, 159
Lilly, S. J., Le Fèvre, O., Renzini, A., et al. 2007, ApJS, 172, 70
Lin, L., Dickinson, M., Jian, H.-Y., et al. 2012, ApJ, 756, 71
Macciò, A. V., Dutton, A. A., van den Bosch, F. C., et al. 2007, MNRAS, 378, 55
Madau, P., & Dickinson, M. 2014, ARA&A, 52, 415
Mandelbaum, R., Seljak, U., Kauffmann, G., Hirata, C. M., & Brinkmann, J. 2006, MNRAS, 368, 715
Mandelbaum, R., Slosar, A., Baldauf, T., et al. 2013, MNRAS, 432, 1544
Mandelbaum, R., Wang, W., Zu, Y., et al. 2016, MNRAS, 457, 3200
Maraston, C. 2005, MNRAS, 362, 799
Maraston, C., Daddi, E., Renzini, A., et al. 2006, ApJ, 652, 85
Martinez-Manso, J., Gonzalez, A. H., Ashby, M. L. N., et al. 2015, MNRAS, 446, 169
Masaki, S., Lin, Y.-T., & Yoshida, N. 2013, MNRAS, 436, 2286
McCracken, H. J., Capak, P., Salvato, M., et al. 2010, ApJ, 708, 202
McCracken, H. J., Milvang-Jensen, B., Dunlop, J., et al. 2012, A&A, 544, A156
McCracken, H. J., Wolk, M., Colombi, S., et al. 2015, MNRAS, 449, 901
Meneux, B., Guzzo, L., de la Torre, S., et al. 2009, A&A, 505, 463
Meneux, B., Guzzo, L., Garilli, B., et al. 2008, A&A, 478, 299
Mo, H. J., & White, S. D. M. 1996, MNRAS, 282, 347
Mo, H. J., & White, S. D. M. 2002, MNRAS, 336, 112
More, S. 2013, ApJL, 777, L26
More, S., Miyatake, H., Mandelbaum, R., et al. 2015, ApJ, 806, 2
More, S., van den Bosch, F. C., Cacciato, M., et al. 2009, MNRAS, 392, 801
More, S., van den Bosch, F. C., Cacciato, M., et al. 2011, MNRAS, 410, 210
Mostek, N., Coil, A. L., Cooper, M., et al. 2013, ApJ, 767, 89
Moster, B. P., Naab, T., & White, S. D. M. 2013, MNRAS, 428, 3121
Moster, B. P., Somerville, R. S., Maulbetsch, C., et al. 2010, ApJ, 710, 903
Nakajima, A., Shioya, Y., Nagao, T., et al. 2008, PASJ, 60, 1249
Navarro, J. F., Frenk, C. S., & White, S. D. M. 1997, ApJ, 490, 493
Noeske, K. G., Weiner, B. J., Faber, S. M., et al. 2007, ApJL, 660, L43
Norberg, P., Baugh, C. M., Gaztañaga, E., & Croton, D. J. 2009, MNRAS, 396, 19
Norberg, P., Baugh, C. M., Hawkins, E., et al. 2001, MNRAS, 328, 64
Norberg, P., Baugh, C. M., Hawkins, E., et al. 2002, MNRAS, 332, 827
Okumura, T., Hikage, C., Totani, T., et al. 2016, PASJ, 68, 38
Ouchi, M., Shimasaku, K., Okamura, S., et al. 2004, ApJ, 611, 685
Pannella, M., Elbaz, D., Daddi, E., et al. 2015, ApJ, 807, 141
Park, C., Vogeley, M. S., Geller, M. J., & Huchra, J. P. 1994, ApJ, 431, 569
Peacock, J. A., & Smith, R. E. 2000, MNRAS, 318, 1144
Press, W. H., & Schechter, P. 1974, ApJ, 187, 425
Reddick, R. M., Wechsler, R. H., Tinker, J. L., & Behroozi, P. S. 2013, ApJ, 771, 30
Rees, M. J., & Ostriker, J. P. 1977, MNRAS, 179, 541
Roche, N., Eales, S. A., Hippelein, H., & Willott, C. J. 1999, MNRAS, 306, 538
Rodighiero, G., Daddi, E., Baronchelli, I., et al. 2011, ApJL, 739, L40
Rubin, V. C., Ford, W. K. J., & Thonnard, N. 1980, ApJ, 238, 471
Rubin, V. C., Thonnard, N. T., & Ford, W. K., Jr. 1982, AJ, 87, 477
Rykoff, E. S., McKay, T. A., Becker, M. R., et al. 2008, ApJ, 675, 1106
Salmon, B., Papovich, C., Finkelstein, S. L., et al. 2015, ApJ, 799, 183
Schreiber, C., Pannella, M., Elbaz, D., et al. 2015, A&A, 575, A74
Scoccimarro, R., Sheth, R. K., Hui, L., & Jain, B. 2001, ApJ, 546, 20
Seljak, U. 2000, MNRAS, 318, 203
Shioya, Y., Taniguchi, Y., Sasaki, S. S., et al. 2008, ApJS, 175, 128
Silverman, J. D., Kashino, D., Sanders, D., et al. 2015, ApJS, 220, 12
Smith, R. E., Peacock, J. A., Jenkins, A., et al. 2003, MNRAS, 341, 1311
Sobral, D., Best, P. N., Geach, J. E., et al. 2010, MNRAS, 404, 1551
Springel, V., Frenk, C. S., & White, S. D. M. 2006, Natur, 440, 1137
Takada, M., Ellis, R. S., Chiba, M., et al. 2014, PASJ, 66, R1
Tamura, N., Takato, N., Shimono, A., et al. 2016, Proc. SPIE, 9908, 99081M
Tinker, J. L., Robertson, B. E., Kravtsov, A. V., et al. 2010, ApJ, 724, 878
Tinker, J. L., Sheldon, E. S., Wechsler, R. H., et al. 2012, ApJ, 745, 16
Tonegawa, M., Totani, T., Okada, H., et al. 2015, PASJ, 67, 81
Totsuji, H., & Kihara, T. 1969, PASJ, 21, 221
Vale, A., & Ostriker, J. P. 2004, MNRAS, 353, 189
van den Bosch, F. C., More, S., Cacciato, M., Mo, H., & Yang, X. 2013, MNRAS, 430, 725
van Uitert, E., Cacciato, M., Hoekstra, H., et al. 2016, MNRAS, 459, 3251

27

Wake, D. A., Whitaker, K. E., Labbé, I., et al. 2011, ApJ, 728, 46
Whitaker, K. E., van Dokkum, P. G., Brammer, G., & Franx, M. 2012, ApJL, 754, L29
White, S. D. M., & Rees, M. J. 1978, MNRAS, 183, 341
Yang, X., Mo, H. J., & van den Bosch, F. C. 2003, MNRAS, 339, 1057
Yang, X., Mo, H. J., & van den Bosch, F. C. 2008, ApJ, 676, 248
Yang, X., Mo, H. J., & van den Bosch, F. C. 2009, ApJ, 695, 900
York, D. G., Adelman, J., Anderson, J. E., Jr., et al. 2000, AJ, 120, 1579

Zahid, H. J., Kashino, D., Silverman, J. D., et al. 2014, ApJ, 792, 75
Zehavi, I., Blanton, M. R., Frieman, J. A., et al. 2002, ApJ, 571, 172
Zehavi, I., Weinberg, D. H., Zheng, Z., et al. 2004, ApJ, 608, 16
Zehavi, I., Zheng, Z., Weinberg, D. H., et al. 2005, ApJ, 630, 1
Zehavi, I., Zheng, Z., Weinberg, D. H., et al. 2011, ApJ, 736, 59
Zhao, D. H., Jing, Y. P., Mo, H. J., & Börner, G. 2009, ApJ, 707, 354
Zheng, Z., Berlind, A. A., Weinberg, D. H., et al. 2005, ApJ, 633, 791
Zheng, Z., Coil, A. L., & Zehavi, I. 2007, ApJ, 667, 760