

UC Berkeley

UC Berkeley Previously Published Works

Title

Target Selection and Validation of DESI Emission Line Galaxies

Permalink

<https://escholarship.org/uc/item/3jd3278g>

Journal

The Astronomical Journal, 165(3)

ISSN

0004-6256

Authors

Raichoor, A

Moustakas, J

Newman, Jeffrey A

et al.

Publication Date

2023-03-01

DOI

10.3847/1538-3881/acb213

Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at <https://creativecommons.org/licenses/by/4.0/>

Peer reviewed



Target Selection and Validation of DESI Emission Line Galaxies

A. Raichoor¹, J. Moustakas², Jeffrey A. Newman³, T. Karim⁴, S. Ahlen⁵, Shadab Alam⁶, S. Bailey¹, D. Brooks⁷, K. Dawson⁸, A. de la Macorra⁹, A. de Mattia¹⁰, A. Dey¹¹, Biprateep Dey³, G. Dhungana¹², S. Eftekharzadeh¹³, D. J. Eisenstein⁴, K. Fanning^{14,15}, A. Font-Ribera¹⁶, J. García-Bellido¹⁷, E. Gaztañaga^{18,19}, S. Gontcho A Gontcho¹, J. Guy¹, K. Honscheid^{14,15}, M. Ishak²⁰, R. Kehoe¹², T. Kisner¹, Anthony Kremin¹, Ting-Wen Lan²¹, M. Landriau¹, L. Le Guillou²², Michael E. Levi¹, C. Magneville¹⁰, M. Manera^{16,23}, P. Martini^{15,24}, Aaron M. Meisner¹¹, Adam D. Myers²⁵, Jundan Nie²⁶, N. Palanque-Delabrouille^{1,10}, W. J. Percival^{27,28,29}, C. Poppett^{1,30,31}, F. Prada³², A. J. Ross¹⁵, V. Ruhlmann-Kleider¹⁰, C. G. Sabiu³³, E. F. Schlafly³⁴, D. Schlegel¹, Gregory Tarlé³⁵, B. A. Weaver¹¹, Christophe Yèche¹⁰, Rongpu Zhou¹, Zhimin Zhou²⁶, and H. Zou²⁶

¹Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Berkeley, CA 94720, USA; araichoor@lbl.gov

²Department of Physics and Astronomy, Siena College, 515 Loudon Road, Loudonville, NY 12211, USA

³Department of Physics & Astronomy and Pittsburgh Particle Physics, Astrophysics, and Cosmology Center (PITT PACC), University of Pittsburgh, 3941 O'Hara Street, Pittsburgh, PA 15260, USA

⁴Center for Astrophysics | Harvard & Smithsonian, 60 Garden Street, Cambridge, MA 02138, USA

⁵Physics Dept., Boston University, 590 Commonwealth Avenue, Boston, MA 02215, USA

⁶Institute for Astronomy, University of Edinburgh, Royal Observatory, Blackford Hill, Edinburgh, EH9 3HJ, UK

⁷Department of Physics & Astronomy, University College London, Gower Street, London, WC1E 6BT, UK

⁸Department of Physics and Astronomy, The University of Utah, 115 South 1400 East, Salt Lake City, UT 84112, USA

⁹Instituto de Física, Universidad Nacional Autónoma de México, Cd. de México C.P. 04510, México

¹⁰IRFU, CEA, Université Paris-Saclay, F-91191 Gif-sur-Yvette, France

¹¹NSF's NOIRLab, 950 N. Cherry Ave., Tucson, AZ 85719, USA

¹²Department of Physics, Southern Methodist University, 3215 Daniel Avenue, Dallas, TX 75275, USA

¹³Universities Space Research Association, NASA Ames Research Centre, USA

¹⁴Department of Physics, The Ohio State University, 191 West Woodruff Avenue, Columbus, OH 43210, USA

¹⁵Center for Cosmology and AstroParticle Physics, The Ohio State University, 191 West Woodruff Avenue, Columbus, OH 43210, USA

¹⁶Institut de Física d'Altes Energies (IFAE), The Barcelona Institute of Science and Technology, Campus UAB, E-08193 Bellaterra, Barcelona, Spain

¹⁷Instituto de Física Teórica (IFT) UAM/CSIC, Universidad Autónoma de Madrid, Cantoblanco, E-28049, Madrid, Spain

¹⁸Institut d'Estudis Espacials de Catalunya (IEEC), E-08034 Barcelona, Spain

¹⁹Institute of Space Sciences, ICE-CSIC, Campus UAB, Carrer de Can Magrans s/n, E-08913 Bellaterra, Barcelona, Spain

²⁰Department of Physics, The University of Texas at Dallas, Richardson, TX 75080, USA

²¹Graduate Institute of Astrophysics and Department of Physics, National Taiwan University, No. 1, Sec. 4, Roosevelt Road, Taipei 10617, Taiwan

²²Sorbonne Université, CNRS/IN2P3, Laboratoire de Physique Nucléaire et de Hautes Energies (LPNHE), F-75005 Paris, France

²³ISera Hunter Fellow, Departament de Física, Universitat Autònoma de Barcelona, Bellaterra, Spain

²⁴Department of Astronomy, The Ohio State University, 4055 McPherson Laboratory, 140 W 18th Avenue, Columbus, OH 43210, USA

²⁵Department of Physics & Astronomy, University of Wyoming, 1000 E. University, Dept. 3905, Laramie, WY 82071, USA

²⁶National Astronomical Observatories, Chinese Academy of Sciences, A20 Datun Road, Chaoyang District, Beijing, 100012, People's Republic of China

²⁷Department of Physics and Astronomy, University of Waterloo, 200 University Avenue W, Waterloo, ON N2L 3G1, Canada

²⁸Perimeter Institute for Theoretical Physics, 31 Caroline Street North, Waterloo, ON N2L 2Y5, Canada

²⁹Waterloo Centre for Astrophysics, University of Waterloo, 200 University Avenue W, Waterloo, ON N2L 3G1, Canada

³⁰Space Sciences Laboratory, University of California, Berkeley, 7 Gauss Way, Berkeley, CA 94720, USA

³¹University of California, Berkeley, 110 Sproul Hall #5800 Berkeley, CA 94720, USA

³²Instituto de Astrofísica de Andalucía (CSIC), Glorieta de la Astronomía, s/n, E-18008 Granada, Spain

³³Natural Science Research Institute, University of Seoul, 163 Seoulsiripdae-ro, Dongdaemun-gu, Seoul, Republic of Korea

³⁴Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, MD 21218, USA

³⁵University of Michigan, Ann Arbor, MI 48109, USA

Received 2022 August 27; revised 2022 December 21; accepted 2023 January 9; published 2023 February 23

Abstract

The Dark Energy Spectroscopic Instrument (DESI) will precisely constrain cosmic expansion and the growth of structure by collecting ~ 40 million extragalactic redshifts across $\sim 80\%$ of cosmic history and one-third of the sky. The Emission Line galaxy (ELG) sample, which will comprise about one-third of all DESI tracers, will be used to probe the universe over the $0.6 < z < 1.6$ range, including the $1.1 < z < 1.6$ range, which is expected to provide the tightest constraints. We present the target selection for the DESI Survey Validation (SV) and Main Survey ELG samples, which relies on the imaging of the Legacy Surveys. The Main ELG selection consists of a g -band magnitude cut and a $(g-r)$ versus $(r-z)$ color box, while the SV selection explores extensions of the Main selection boundaries. The Main ELG sample is composed of two disjoint subsamples, which have target densities of about 1940 deg^{-2} and 460 deg^{-2} , respectively. We first characterize their photometric properties and density variations across the footprint. We then analyze the DESI spectroscopic data that have been obtained from 2020 December to 2021 December in the SV and Main Survey. We establish a preliminary criterion for selecting reliable redshifts, based on the [O II] flux measurement, and assess its performance. Using this criterion, we are able to



Original content from this work may be used under the terms of the [Creative Commons Attribution 4.0 licence](https://creativecommons.org/licenses/by/4.0/). Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

present the spectroscopic efficiency of the Main ELG selection, along with its redshift distribution. We thus demonstrate that the Main selection 1940 deg^{-2} subsample alone should provide 400 deg^{-2} and 460 deg^{-2} reliable redshifts in the $0.6 < z < 1.1$ and the $1.1 < z < 1.6$ ranges, respectively.

Unified Astronomy Thesaurus concepts: [Emission line galaxies \(459\)](#); [Surveys \(1671\)](#); [Large-scale structure of the universe \(902\)](#)

1. Introduction

Since the observation of the acceleration of the expansion of the universe (Riess et al. 1998; Perlmutter et al. 1999), the cosmology community has focused its efforts on gathering the data to provide more precise potential observational constraints. Several cosmological probes have been used in order to produce independent measurements with different systematics (see Weinberg et al. 2013 for a review), the most established methods being Type Ia supernovae and baryonic acoustic oscillations (BAO), to constrain the geometry of the universe, and weak lensing, galaxy clusters, and redshift-space distortions (RSD), to constrain the growth of structure. To reach the goal, dedicated facilities survey large fractions of the sky, with high-quality imaging (e.g., the Dark Energy Survey, or DES: the Dark Energy Survey Collaboration 2005; Hyper Suprime-Cam, or HSC: Aihara et al. 2018; Euclid: Laureijs et al. 2011; LSST: Ivezić et al. 2019) and/or massive spectroscopy (e.g., Two-Degree Field Galaxy Redshift Survey, or 2dFGRS: Colless et al. 2003; Six-Degree Field Galaxy Redshift Survey, or 6dFGRS: Jones et al. 2009; BOSS: Dawson et al. 2013; WiggleZ: Drinkwater et al. 2010; eBOSS: Dawson et al. 2016; the Dark Energy Spectroscopic Instrument, or DESI, Euclid, and the Subaru Prime Focus Spectrograph, or PFS: Takada et al. 2014).

Massive spectroscopic surveys probe the large-scale structure (LSS) of the matter distribution, by measuring the spectroscopic redshifts (z_{spec}) of a vast number of galaxies over large areas and different epochs. One strength of this approach is that the same data set allows one to constrain, at the same time, the geometry of the universe, with the BAO scale (Eisenstein & Hu 1998), and the growth of structure, with the RSD (Kaiser 1987) method. The Sloan Digital Sky Survey (SDSS) experiment (York et al. 2000) has been a pioneer of such surveys, with the co-first BAO measurement (Eisenstein et al. 2005). Alam et al. (2021) have summarized and analyzed 20 yr of SDSS spectroscopic observations of about two million z_{spec} over $0 < z < 5$ and $10,000 \text{ deg}^2$, which led to state-of-the-art constraints on the Hubble constant ($H_0 = 68.18 \pm 0.79 \text{ km s}^{-1} \text{ Mpc}^{-1}$) and the σ_8 parameter normalizing the growth of structure ($\sigma_8 = 0.85 \pm 0.03$).

The DESI experiment (Levi et al. 2013; DESI Collaboration et al. 2016a, 2016b) will pursue this effort and increase the number of observed z_{spec} by an order of magnitude, with about 40 million extragalactic z_{spec} over $14,000 \text{ deg}^2$. DESI will follow the same approach as SDSS, by using an optimized tracer for each targeted redshift range. About 13 million galaxies from the bright galaxy sample (BGS) will cover the $0.05 < z < 0.4$ range, about eight million luminous red galaxies (LRGs) will cover the $0.4 < z < 1.1$ range, about 16 million emission line galaxies (ELGs) will cover the $0.6 < z < 1.6$ range, and, last, about three million quasars (QSOs) will cover the $z > 0.9$ range, be used as tracers in the $0.9 < z < 2.1$ range, and use Ly- α forests as a probe of the intergalactic medium at $z > 2.1$. Additionally, DESI will also observe about 10 million stars from the Milky Way Survey (MWS). The BGS and MWS

programs will be observed in “bright” time, i.e., when the Moon is up, whereas the other tracers (LRGs, ELGs, and QSOs) will be observed in “dark” time, i.e., when the Moon is down.

This paper is dedicated to the DESI ELG sample, which is composed of star-forming galaxies. The goal of the ELG sample is to take advantage of the following two facts: (1) the universe’s star formation rate density peaks at $z \sim 1-2$ (e.g., Madau & Dickinson 2014), thus star-forming galaxies are very common at that epoch; and (2) the ELG z_{spec} can be reliably measured in a rather short period of observation time, as it only requires a significant detection of the emission lines in the spectrum, with no need to detect the continuum in a significant way—in particular, the [O II] doublet $\lambda\lambda 3726, 29 \text{ \AA}$ offers an unambiguous signature of the z_{spec} (see, for instance, Moustakas et al. 2006 for the link between the [O II] line strength and the star formation). Some reference spectroscopic surveys have sampled the ELG population at $z \sim 1-2$ over a few square degrees (e.g., VVDS: Le Fèvre et al. 2013; DEEP2: Newman et al. 2013), paving the way for their use in spectroscopic cosmological experiments. For the above reasons, the ELG tracer is a key tracer in this decade of massive spectroscopic surveys (e.g., DESI, Euclid, and PFS), and it will constitute about one-third of the DESI spectra. The DESI ELG sample will probe the universe over the $0.6 < z < 1.6$ range, and in particular over the $1.1 < z < 1.6$ range, which will result in the tightest DESI cosmological constraints. It will be the very first survey to densely sample this redshift range, providing faint targets that have not been extensively explored by any previous survey. For instance, the eBOSS ELG sample (Raichoor et al. 2017, 2021) has a target density about 10 times smaller and targets about one magnitude brighter than the DESI ELG sample.

In this respect, the DESI ELG sample will be the first of its kind, which thus involves several challenges. First, the target density needs to be high, about 2400 deg^{-2} , because the ELG targets will be assigned fibers after the LRG and QSO targets, so the selection must provide enough targets for each fiber to be able to reach a target most of the time. This requires the selection of rather faint targets; however, another constraint is the requirement for the number of z_{spec} measurement failures in a typical DESI exposure (15 minutes, in nominal conditions) to remain a reasonable fraction of the observed spectra. For this purpose, a large enough fraction of the targets need to have sufficient [O II] fluxes to secure a reliable z_{spec} measurement. A quantified requirement is that the DESI ELG target sample should provide at least 400 deg^{-2} reliable z_{spec} in both the $0.6 < z < 1.1$ and $1.1 < z < 1.6$ ranges, as Fisher forecasts show that this will be sufficient to reach the DESI experiment’s required cosmological precision (DESI Collaboration et al. 2023a, in preparation). Last, as for other tracers, the DESI ELG sample must have a fraction of catastrophic z_{spec} measurements (“catastrophics”) as low as possible (of the order of 1%), as the LSS analysis is very sensitive to catastrophic z_{spec} .

To meet these requirements, the DESI experiment conducted two programs (2020 December–2021 May), before starting the actual Main Survey in 2021 May. The first program was the Survey Validation (SV)³⁶ program (2020 December–2021 March), which consisted of deep observations of extended target selections for all tracers. Those SV data have been used to fine-tune the Main Survey target selections. The second program was the One-Percent Survey, hereafter called “One-Percent”³⁷ (2021 April–May), where the target selections that were close or equal to the Main Survey ones were observed at a very high completeness, over 140 deg².

This paper is part of a series of papers presenting the DESI target selections and their characterization. The DESI Collaboration et al. (2023a, in preparation) present an overview of the DESI spectroscopic observations and tracers used by those papers. Myers et al. (2022) show how those target selections are implemented in DESI. Lan et al. (2022) and Alexander et al. (2022) show the construction of spectroscopic truth tables, based on visual inspections (VIs), for the galaxies (BGS, LRG, and ELG) and the QSO targets, respectively. The MWS sample is presented in Cooper et al. (2022), the BGS sample is presented in Hahn et al. (2022), the LRG sample is presented in Zhou et al. (2022), the ELG sample is presented in this paper, and the QSO sample is presented in Chaussidon et al. (2022). These five target selection papers present the final DESI samples, and supersede the preliminary target selections that were presented in Allende Prieto et al. (2020), Ruiz-Macias et al. (2020), Zhou et al. (2020), Raichoor et al. (2020), and Yèche et al. (2020).

This paper is structured as follows. Section 2 presents the imaging, the footprints, and the photometry used to select the ELG targets. We then present the Main Survey ELG target selection in Section 3, the SV ELG target selection in Section 4, and the Main Survey ELG sample photometric properties in Section 5. Section 6 introduces the DESI spectroscopic data (The SV, One-Percent, and Main Survey observations up to 2021 December), which are used in Section 7 to analyze the spectroscopic properties of the Main Survey ELG sample. We conclude in Section 8.

All magnitudes are in the AB system (Oke & Gunn 1983), and corrected for Galactic extinction using the Schlegel et al. (1998) maps. All displayed sky maps use the HEALPix scheme (Górski et al. 2005), with a resolution of 0.21 deg² ($n_{\text{side}} = 128$), but the computation in Section 5.3 uses a finer resolution of 0.05 deg² ($n_{\text{side}} = 256$).

2. Imaging, Footprints, and Photometry

The DESI ELG targets are selected from the grz photometry of Data Release 9 of the Legacy Imaging Surveys³⁸ (LS-DR9; D. J. Schlegel et al. 2023, in preparation). This release covers about 19,700 deg² in the optical grz bands, complemented with Wide-field Infrared Survey Explorer near-IR data (Meisner et al. 2021). We present here a brief description of the optical imaging and photometry, focusing on the parts that are relevant for the ELG target selection, and we refer the reader to D. J. Schlegel et al. (2023, in preparation) for more details.

³⁶ This is labeled “SV1” in the DESI files conventions, and also in some other DESI papers.

³⁷ This is labeled “SV3” in the DESI files conventions, and also in some other DESI papers.

³⁸ <https://www.legacysurvey.org/dr9>

2.1. Imaging

The optical grz imaging comes from several observing programs. For the northern part of the North Galactic Cap (NGC), the imaging comes from two programs: the Beijing–Arizona Sky Survey (BASS; Zou et al. 2017) provides the g and r bands, observed with the 90Prime camera on the Bok 2.3 m telescope; and the Mayall z -band Legacy Survey (MzLS) provides the z band, observed with the Mosaic-3 camera on the 4 m Mayall telescope at Kitt Peak National Observatory. For the southern part of the NGC and the South Galactic Cap (SGC), the imaging mostly comes from two programs: the Dark Energy Camera Legacy Survey (DECaLS; Dey et al. 2019) and DES (The Dark Energy Survey Collaboration 2005)—both use the Dark Energy Camera (Flaugher et al. 2015) on the 4 m Blanco telescope at the Cerro Tololo Inter-American Observatory.

We note that the DECaLS, BASS, and MzLS surveys followed a dynamic observing strategy, to achieve, as far as possible, a uniform depth across the footprints. In particular, the considered depths account for the Galactic extinction—i.e., the imaging is deeper in regions with high Galactic extinction—so the target selection should be less sensitive to the Galactic dust map (see Section 6.2 of Dey et al. 2019). Nevertheless, because of the capping of the individual imaging exposure times, this strategy cannot be applied to the g -band imaging for $E(B - V) \gtrsim 0.15$, as the g band has the largest extinction factor.³⁹ DES did not follow such a strategy, but as it was fairly deep, the effect of the Galactic extinction on the imaging depth was less critical for the ELG targets.

Figure 1 illustrates this approach for the g band, where the extinction factor is the largest. In regions of high extinction (top), the g -band depth (middle) is larger, resulting in a rather homogeneous extinction-corrected depth map in each of the three footprints of the imaging program (bottom).

2.2. Footprints

As a result of different programs providing the imaging, different parts of the footprint have different imaging depths, which matters for the ELG target selection, as these tracers are faint in imaging. To illustrate this point, Figure 2 displays normalized cumulative distributions of the signal-to-noise ratio (S/N) in the selection band⁴⁰ over the nominal DESI footprints for its three dark tracers. The ELG targets have a typical S/N of 13 in the imaging, whereas the QSO and LRG targets have a typical S/Ns of 36 and 60, respectively, and thus are less sensitive to depth variations across the footprints.

For this reason, we define three footprints, which will be analyzed separately in this paper: the North, corresponding to the decl. $> 32^{\circ}.375$ part of the NGC, covered by BASS and MzLS; the South-DECaLS, corresponding to the non-DES SGC part and the decl. $< 32^{\circ}.375$ part of the NGC; and the South-DES, corresponding to the DES imaging in the SGC. These footprints can be visualized in the depth maps in Figure 1, where the North is displayed in blue–green, the South-DECaLS in yellow–orange, and the South-DES in red. Table 1 reports the approximate areas and imaging depths per footprint. One notices that the North footprint is about 0.5 mag

³⁹ The $A/E(B - V)$ coefficients are 3.214, 2.165, and 1.211 for the g , r , and z band, respectively; see <https://www.legacysurvey.org/dr9/catalogs/#galactic-extinction-coefficients>.

⁴⁰ For instance, $S/N = \text{flux}_g \times \sqrt{\text{flux_ivar}_g}$, for the ELG targets.

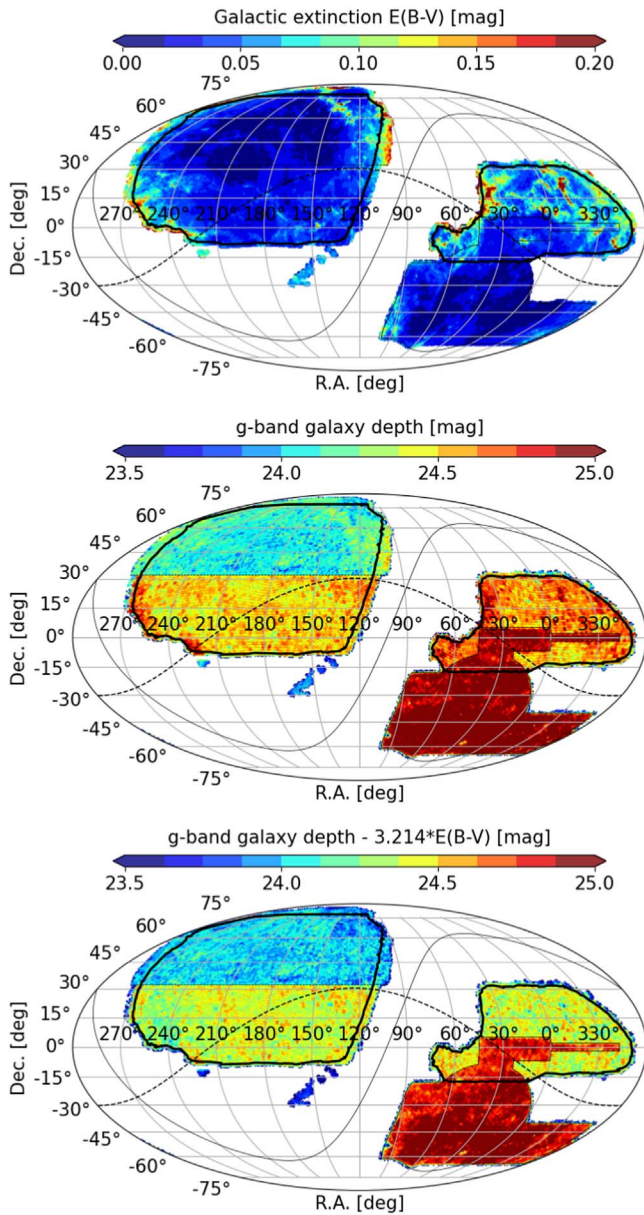


Figure 1. Sky maps of the Galactic dust extinction (top), the imaging g -band depth (middle), and the extinction-corrected g -band depth (bottom), for the imaging data used to select the DESI ELG targets. The depths are 5σ depths for a $0''.45$ radius exponential profile, typical of DESI ELG targets. The thick black line represents the $14,000 \text{ deg}^2$ footprint covered by DESI. The Galactic plane is displayed as a solid line and the Sagittarius plane is displayed as a dashed line.

shallower in the g and r bands than the South-DECaLS footprint, and that the South-DES footprint is about 0.5 – 1.0 mag deeper than the South-DECaLS footprint in all three grz bands. As ELG targets are faint, those depth differences impact the target selection, in terms of the detected objects and contamination, as will be seen later in the paper.

2.3. Photometry

The overall data reduction and photometry is performed with the `legacypipe`⁴¹ pipeline. The LS-DR9 images are astrometrically calibrated with Gaia Data Release 2 (DR2;

⁴¹ <https://github.com/legacysurvey/legacypipe>

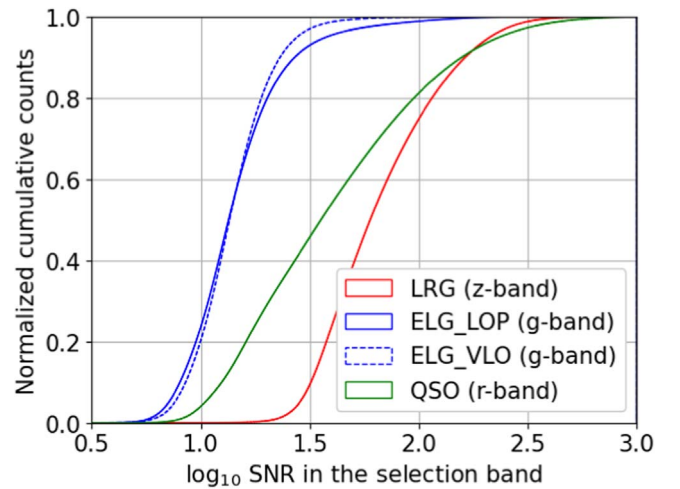


Figure 2. Cumulative distribution of the logarithm of the photometric S/N (see the text) in the selection band for the DESI Main Survey dark tracers over the nominal DESI footprints. For the ELG targets, the two Main Survey selections are shown.

Gaia Collaboration et al. 2018) and photometrically calibrated with Pan-STARRS 1 (Chambers et al. 2016), using color terms to place the photometry on the same system as the LS-DR9 one. The photometry is performed with the `Tractor` software (Lang 2016; D. Lang et al. 2023, in preparation). Source detection is conducted on stacked images, then all measurements are based on individual exposures. Each source is modeled with an analytic profile (point-source, exponential with fixed parameters, exponential, de Vaucouleurs, or Sérsic), and a model image is generated for each exposure. Increasingly more complex profiles are allowed for sources detected with higher S/Ns. The source properties (position, shape, and flux) are measured through a likelihood optimization (χ^2 minimization) of the set of model images covering the considered region.

Based on the best-fit properties, `Tractor` provides the total flux of each source, as well as its “fiber flux,” which corresponds to the predicted flux within a fiber of diameter $1''.5$ —the size of a DESI fiber—for a $1''$ Gaussian seeing. Those fiber fluxes therefore predict the typical amount of light that a DESI fiber would see.

3. Main Survey Target Selection

In this section, we present the DESI Main Survey ELG target selection. The selection cuts are detailed in Table 2 and illustrated in Figure 3. The extended selection explored in the DESI SV program, which was used to finalized those Main Survey cuts, is presented in Section 4.

The DESI ELG sample is the first of its kind, with no significant previous reference sample having been observed so far. For instance, the VVDS and DEEP2 observations covered a few square degrees, while the WiggleZ (Drinkwater et al. 2010) and eBOSS/ELG (Raichoor et al. 2017, 2021) surveys observed about 1000 square degrees, but their ELG samples were more than one magnitude brighter, had a lower density by five to 10 times (about 200 – 400 deg^{-2}), and only extended to $z < 1.1$. The first proposed DESI ELG selection was based on a simple $g - r$ versus $r - z$ selection (DESI Collaboration et al. 2016a), although it could not be spectroscopically tested at that time. Karim et al. (2020) explored that selection, along with more advanced ones, with dedicated spectroscopic

Table 1
Imaging Properties and ELG Target Density per Footprint for Each of the Two Main ELG Samples (ELG_LOP and ELG_VLO)

Footprint	LS-DR9 Area (deg ²)	DESI Area (deg ²)	<i>g</i> -depth (AB mag)	<i>r</i> -depth (AB mag)	<i>z</i> -depth (AB mag)	ELG_LOP density (deg ⁻²)	ELG_VLO density (deg ⁻²)
North	5100	4400	24.1	23.5	23.0	1930	410
South-DECaLS	9500	8500	24.5	23.9	23.0	1950	490
South-DES	5100	1100	24.9	24.7	23.5	1900	480

Note. The areas are approximate. The depths are 5σ depths for a $0''.45$ radius exponential profile, typical of DESI ELGs.

observations. This pilot program demonstrated that all the tested selections had similar overall performances.

Based on that analysis, for the sake of simplicity and robustness, we choose simple color–color cuts in the $g - r$ versus $r - z$ diagram to select the DESI ELG targets. Hereafter, we describe the chosen cuts.

3.1. The “ELG_LOP” and “ELG_VLO” Subsamples

As mentioned in Section 1, the goal of the DESI ELG sample is to provide cosmological constraints over the $0.6 < z < 1.6$ range, favoring the $1.1 < z < 1.6$ range as far as possible, where other DESI samples are the least dense. To do so, the ELG sample is split in two disjoint samples: “ELG_LOP” at ~ 1940 deg⁻² and “ELG_VLO” at ~ 460 deg⁻².⁴² We remind the reader that the DESI observations use priorities to assign fibers to targets (A. Raichoor et al. 2023, in preparation). ELG_LOP has higher priority in the fiber assignment, and favors the $1.1 < z < 1.6$ range, whereas ELG_VLO has lower priority, and favors the $0.6 < z < 1.1$ range. With this fiber assignment configuration, cosmological Fisher forecasts demonstrate that the ELG sample fulfills the expected performance (DESI Collaboration et al. 2023a, in preparation).

The names of these two samples—ELG_LOP and ELG_VLO—are names that are assigned to targeting bits by `desitarget`, the target selection pipeline (Myers et al. 2022), and indicate the ELG priority states in the fiber assignment (“low” and “very low”).

3.2. “ELG_HIP” Subsample

For the dark tiles, the tracers in order of decreasing fiber assignment priority are: QSOs, LRGs, ELG_LOP, and ELG_VLO. This results in very high fiber assignment rates for the QSO and LRG targets, but lower ones for the ELG_LOP targets, and even lower ones for the ELG_VLO targets. In order to have a significant number of observed pairs of ELG and LRG targets, a third ELG sample is defined, ELG_HIP, which is a 10% random subsampling of the ELG_LOP and ELG_VLO samples, but with the same fiber assignment priority as one of the LRG targets. This provides more information about the small-scale cross-correlation between the ELG targets and higher-priority targets. Without this extra sample, the small-scale effects of fiber collisions, together with the preference for always observing the higher-priority objects, would significantly increase the noise for cosmological analyses cross-correlating ELG targets with LRG targets (see e.g., Bianchi et al. 2018 and Mohammad et al. 2020 for the weight computation method).

⁴² The area for computing these densities does not account for the $\sim 1\%$ area removed by the angular masking, described in Section 3.3.1.

Similar to ELG_LOP and ELG_VLO, the ELG_HIP name is assigned to targeting bits by `desitarget`, and indicates the ELG priority state in the fiber assignment (“high”). As this ELG_HIP sample is a random subsample of the ELG_LOP and ELG_VLO samples, we do not discuss it further in this paper.

3.3. Main Survey Selection Cuts

The Main Survey ELG selection cuts are detailed in Table 2 and illustrated in Figure 3. They are of three kinds: (1) quality cuts, to ensure that the photometry is reliable; (2) a cut in the *g*-band fiber magnitude; and (3) a selection box in the $g - r$ versus $r - z$ diagram.

We underline that the cuts are the same in the North and South-DECaLS/DES footprints, even though the photometric systems are slightly different. This choice has been motivated by several reasons: the exact color transformation between the two systems is not trivial, as it depends on the considered object (e.g., star, blue, or red galaxies); the North has different imaging systematics than the South, in particular in the *g*-band depth; and with the partial sampling of the SV program, it was not possible to define a secure tuning of the selection cuts, which would provide a similar ELG redshift distribution in the three footprints, as the redshift distribution has nontrivial dependencies on the imaging and foreground variations. For the sake of simplicity, we thus keep the same cuts in the three footprints.

3.3.1. Quality Cuts

Quality cuts are designed to select sources with reliable photometric measurements. For computation reasons, the `legacypipe` pipeline processes the sky in $0^\circ 25' \times 0^\circ 25'$ bricks, which slightly overlap. The `brick_primary` cut requests the object not to be in the LS-DR9 overlap between two bricks. We further require that there is at least one observation in each of the three *grz* bands, and that the measured flux has a positive S/N in all three bands (i.e., a positive flux and a non-null inverse variance). We also apply a minimal angular masking, to reject regions around very bright stars (Gaia $G < 13$) and large galaxies (J. Moustakas et al. 2023, in preparation) or globular clusters. We emphasize that this masking is purposely minimal, and common to all three DESI dark tracers. Further a posteriori masking will be applied on the spectroscopic data in the analyses, as, for instance, we only know from the target density variations that the ELG target selection has spurious targets around moderately bright stars ($13 < \text{Gaia } G < 16$).

3.3.2. *g*-band Magnitude Cut

We then make a selection on the *g*-band magnitude, which is motivated by the fact that the [O II] flux best correlates with the

Table 2
Main Survey Target Selection Cuts

Sample	Density	Cuts	Comment
Clean	...	brick_primary = True nobs_{grz} > 0 $\text{flux}_{\{grz\}} \times \sqrt{\text{flux_ivar}_{\{grz\}}} > 0$ (maskbits & 2 ¹) = 0, (maskbits & 2 ¹²) = 0, (maskbits & 2 ¹³) = 0	Unique object Observed in the grz bands Positive S/N in the grz bands Not close to bright star/galaxy
ELG_LOP	~1940 deg ⁻²	Clean (g > 20) and (g _{fib} < 24.1) 0.15 < r - z g - r < 0.5 × (r - z) + 0.1 g - r < -1.2 × (r - z) + 1.3	Clean sample Magnitude cut r - z cut Star/low-z cut Redshift/[O II] cut
ELG_VLO	~460 deg ⁻²	Clean (g > 20) and (g _{fib} < 24.1) 0.15 < r - z g - r < 0.5 × (r - z) + 0.1 (g - r > -1.2 × (r - z) + 1.3) and (g - r < -1.2 × (r - z) + 1.6)	Clean sample Magnitude cut r - z cut Star/low-z cut Redshift/[O II] cut

Note. The cuts are the same for the North and the South-DECaLS/DES regions. We use the following definitions: $\{grz\} = 22.5 - 2.5 \cdot \log_{10}(\text{flux}_{\{grz\}}/\text{mw_transmission}_{\{grz\}})$, $g_{\text{fib}} = 22.5 - 2.5 \cdot \log_{10}(\text{fiberflux}_g/\text{mw_transmission}_g)$. The brick_primary, nobs_{grz}, flux_{grz}, fiberflux_g, flux_ivar_{grz}, mw_transmission_{grz}, and maskbits columns are described online: <https://www.legacysurvey.org/dr9/catalogs/>.

bluest band flux (Comparat et al. 2016); this ensures that the selection favors [O II] emitters. We discard bright objects, which are unlikely to be at $z > 0.6$; as these represent a marginal fraction of the ELG sample, it does not matter if we cut on the fiber or total magnitude. The faint cut in the g -band fiber magnitude is tuned to reach the desired densities of about 1940 deg⁻² for ELG_LOP and 460 deg⁻² for ELG_VLO. A cut on the fiber magnitudes was favored over a cut on the total magnitudes, because the latter provide more z_{spec} failures, due to galaxies not having enough flux in the DESI fiber.

3.3.3. $g - r$ versus $r - z$ Selection

The ELG_LOP and ELG_VLO samples rely on simple $g - r$ versus $r - z$ cuts. The primary motivation of the cuts is the redshift selection, as illustrated in Figure 3. That figure displays the density of $g_{\text{fib}} < 24.1$ objects in the LS-DR9 catalogs, where the color-coding indicates the mean photometric redshift (z_{phot}) from HSC/DR2 (Tanaka et al. 2018; Aihara et al. 2019). These z_{phot} measurements are of exquisite quality for our magnitude and redshift ranges of interest, thanks to the depth and wavelength coverage of the HSC data; Karim et al. (2020) have already illustrated this point with previous HSC data, and in Appendix B we further illustrate how the data compare with the DESI ELG z_{spec} .

The slanted cut with a positive slope ($g - r < 0.5 \times (r - z) + 0.1$) in common between the ELG_LOP and ELG_VLO selections rejects stars and galaxies at $z < 0.6$.

The ELG_LOP $r - z > 0.15$ cut rejects $z > 1.6$ galaxies for which the [O II] doublet is outside the DESI spectrograph coverage, so that no reliable z_{spec} can be expected. The ELG_LOP slanted cut with a negative slope ($g - r < -1.2 \times (r - z) + 1.3$) optimizes the fraction of $1.1 < z < 1.6$ targets with high [O II] flux, as this is the goal for this sample. At first order, the redshift is driving this cut, as shown by the HSC z_{phot} measurements. At second order, favoring the [O II] emitters pushes this cut to the blue, as illustrated by the stellar evolution tracks on Figure 3. Those tracks show two simple Bruzual & Charlot (2003) evolution models of galaxies computed with

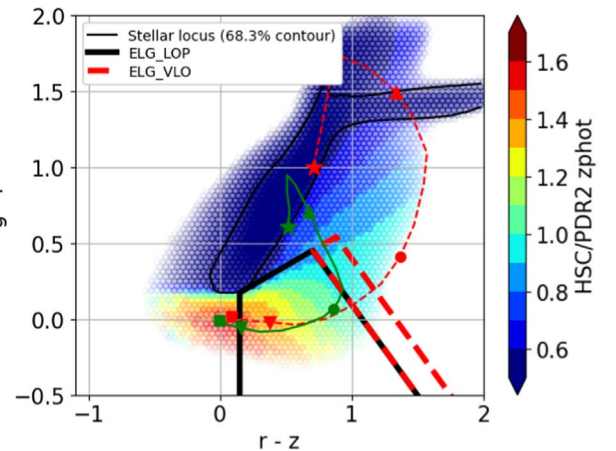


Figure 3. Main ELG target selection cuts in the $g - r$ vs. $r - z$ diagram. The “ELG_LOP” and “ELG_VLO” selections are displayed as the solid black lines and the dashed red lines, respectively. The background symbols describe a $g_{\text{fib}} < 24.1$ sample; the color-coding indicates the mean HSC/DR2 z_{phot} measurements, and the transparency scales with the logarithm of the density. The thin black contour indicates the stellar locus. The tracks show the stellar evolutions of a highly star-forming galaxy (solid green line) and a moderately star-forming galaxy (dashed red line), observed at z_{spec} 2.0, 1.6, 1.1, 0.6, and 0.1 (the square, downward triangle, circle, upward triangle, and star, respectively). See the text for more details.

EzGal (Mancone & Gonzalez 2012). The two galaxy models are formed at $z = 3$, with simple exponentially declining star formation histories (i.e., with a star formation rate that is proportional to $e^{-\text{age}/\tau}$). One is moderately star-forming ($\tau = 1$ Gyr; dashed red line), the other one is more star-forming ($\tau = 5$ Gyr; solid green line); the symbols illustrate where such galaxies would be in the $g - r$ versus $r - z$ diagram at different observation redshifts—as expected, at a fixed redshift, galaxies with bluer colors are more star-forming. For instance, one sees at $z = 1.1$ (circles) that the more star-forming model is about 0.5 mag bluer in $r - z$ than the other one.

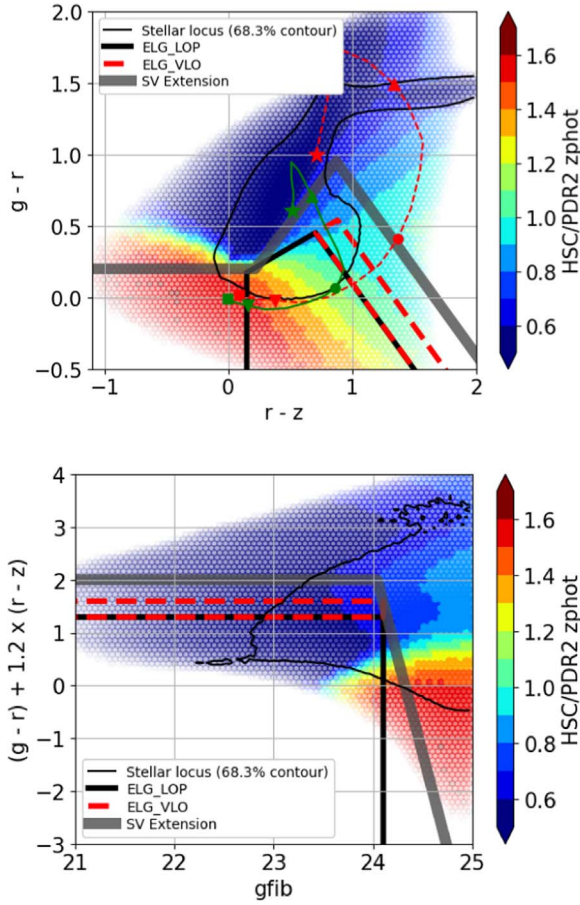


Figure 4. SV ELG target selection cuts in the $g - r$ vs. $r - z$ diagram (top) and $(g - r) + 1.2 \times (r - z)$ vs. g_{fib} diagram (bottom). The ELG_LOP and ELG_VLO selections are displayed as the solid black lines and dashed red lines, respectively. The background symbols describe a $g_{\text{fib}} < 25.0$ sample; the color-coding indicates the mean HSC/DR2 z_{phot} measurements, and the transparency scales with the logarithm of the density. The contours and tracks are as in Figure 3.

The ELG_VLO slanted cuts with a negative slope ($(g - r > -1.2 \times (r - z) + 1.3)$ and $(g - r < -1.2 \times (r - z) + 1.6)$) are an extension of the ELG_LOP selection toward redder colors and hence lower redshifts. The reddest cut is driven to remove $z < 0.6$ galaxies. We remind the reader that the ELG_VLO sample is disjoint from the ELG_LOP sample.

3.3.4. Target Density

The cuts described above provide an ELG_LOP sample of about 1940 deg^{-2} and an ELG_VLO sample of about 460 deg^{-2} . Because of the different imaging properties—in particular, the depths—of the three North, South-DECaLS, and South-DES footprints, the actual average density over each footprint is slightly different, as reported in the last two columns of Table 1: from 1900 deg^{-2} to 1950 deg^{-2} for the ELG_LOP sample, and from 410 deg^{-2} to 490 deg^{-2} for the ELG_VLO sample.

4. SV Target Selection

In this section, we describe the DESI SV target selection, which expands the Main Survey selections described in

Section 3. The SV selection cuts are detailed in Appendix A, and illustrated in Figure 4.

4.1. Motivations

The SV ELG sample has been designed to provide information for finalizing the Main Survey ELG selections. The only existing magnitude-limited spectroscopic reference samples probing the desired DESI ELG magnitudes are limited to a few square degrees (e.g., DEEP2: Newman et al. 2013; VVDS: Le Fèvre et al. 2013). Besides, DESI being a new instrument, its ability to measure reliable z_{spec} for targets as faint as ELG ones needs to be thoroughly tested. For these two reasons, the DESI SV ELG sample explores a rather large photometric space, with a target density of about 7000 deg^{-2} .

4.2. SV Selection Cuts

Hereafter, we detail the philosophy of the DESI SV ELG selection cuts reported in Appendix A and illustrated in Figure 4.

4.2.1. $g - r$ versus $r - z$ Extensions

The first extensions to be explored relax the Main Survey cuts in the $g - r$ versus $r - z$ diagram, as illustrated in the top panel of Figure 4. The cuts are generously extended toward bluer $r - z$ colors, with a $g - r < 0.2$ cut to securely remove low-redshift galaxies and stars. According to HSC z_{phot} measurements, that region should include a significant fraction of redshifts, in the range $1.1 < z < 1.6$, and it has been extremely poorly explored so far. While this region is very valuable for DESI, with $1.1 < z < 1.6$ ELG targets, it is also costly, because any $z > 1.6$ target would not provide a reliable z_{spec} , as the [O II] doublet would be outside the DESI spectrograph coverage.

The cuts are also slightly extended toward low-redshift galaxies and the stellar locus (the positive slope cut). Existing spectroscopic data and HSC z_{phot} measurements consistently show that there is a sharp transition, with a density of $z < 0.6$ objects quickly rising when going to redder $g - r$ colors. As DESI is expected to provide reliable z_{spec} for most of those, there only is a marginal need to explore this region.

Last, the cuts are extended toward redder $r - z$ colors, to cover the eBOSS/ELG selection region. From HSC z_{phot} and eBOSS/ELG z_{spec} measurements, we know that this region mostly hosts $z < 1.1$ galaxies. This extension is motivated by the early desire for an overlap with the eBOSS/ELG sample, and to secure a fallback Main Survey selection, in the unlikely case the DESI instrument were to perform far worse than expected.

4.2.2. Faint Extensions (Sliding Cut)

An important extension explores the faint end of the target selection, to test the ability of the DESI instrument to provide reliable z_{spec} there. Provided that the target density significantly increases when going fainter, this extension is restricted to blue objects—the most interesting ones for ELGs—to prevent the SV sample from being overwhelmed by faint objects. To do so, we adopt a sliding cut in the $g - r$ versus $r - z$ diagram, as illustrated in the bottom panel of Figure 4. The sliding cut uses the $(g - r) + 1.2 \times (r - z)$ color, which broadly scales as the [O II] flux. On the red $r - z$ side, this cut restricts the sample to

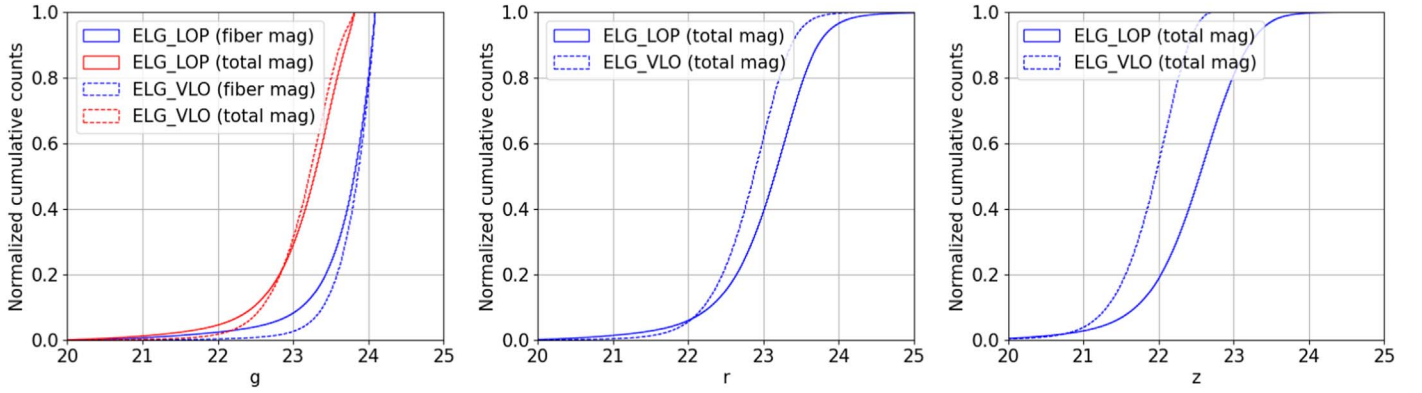


Figure 5. ELG_LOP (solid lines) and ELG_VLO (dashed lines) total target magnitude normalized cumulative distributions in the g band (left), r band (middle), and z band (right). For the g band, we also display the fiber magnitude distributions (red lines).

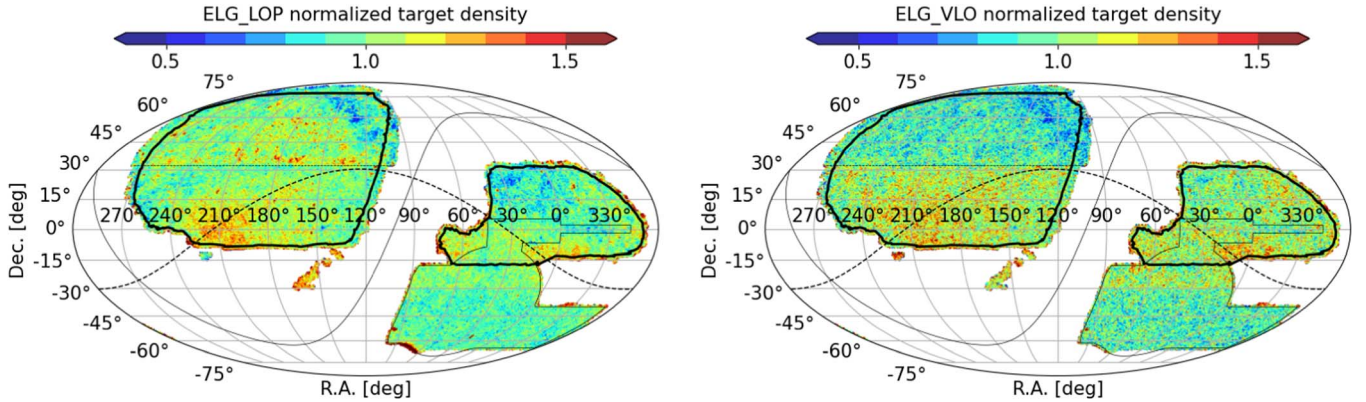


Figure 6. Main Survey ELG_LOP (left) and ELG_VLO (right) sample density sky maps. The density is divided by the overall average value (1940 deg^{-2} for the ELG_LOP sample and 460 deg^{-2} for the ELG_VLO sample), to display the fractional difference to the average. The thick black line represents the $14,000 \text{ deg}^2$ footprint covered by DESI. The Galactic plane is displayed as a solid line, while the Sagittarius plane is displayed as a dashed line.

bright objects, as the faint objects there are expected to have a marginal [O II] flux, and are thus unlikely to provide a reliable z_{spec} . On the blue $r - z$ side, this cut explores targets that are fainter by few tenths of magnitudes, which are expected to have a significant [O II] flux, and hence would provide reliable z_{spec} .

4.2.3. g_{tot} and g_{fib} Extensions

Finally, all the above cuts are applied on samples restricted in g_{tot} , the total g -band magnitude, or in g_{fib} , the fiber g -band magnitude. While a g_{tot} -limited sample corresponds to a better defined galaxy population, it could contain a significant fraction of targets with too small fluxes in the DESI fibers to provide reliable z_{spec} . A g_{fib} -limited sample has the advantage of being more homogeneous and complete in terms of reliable z_{spec} .

5. Photometric Properties of the Main Sample

This section presents a preliminary discussion of the Main Survey ELG sample density fluctuations across the footprint, which are driven by variations in both the LS-DR9 imaging properties and the astrophysical foreground maps (e.g., Galactic stellar density and dust extinction). Given that the ELG target magnitudes are close to the imaging depth, this sample is more sensitive to these imaging/foreground variations than the other DESI dark tracers, and it is likely that significant work dedicated to accounting for them will be

required in the data analysis, which will need to remove such dependencies prior to a cosmological analysis being performed.

Besides, the final LSS ELG sample will be restricted to objects with a reliable redshift in the $0.6 < z < 1.6$ range. Both the ELG target density fluctuations and redshift efficiency variations with spectroscopic observing conditions will have to be corrected to produce reliable cosmological results.

This is why we hereafter restrict ourselves to simple diagnoses, in order to illustrate the overall properties of the Main Survey ELG sample.

5.1. Magnitude Distributions

Figure 5 displays the ELG_LOP (solid lines) and ELG_VLO (dashed lines) normalized cumulative distributions of the target magnitudes. For the g band, we present both the fiber magnitude, used for the selection, and the total magnitude. For the r and z bands, we only present the total magnitude, as these come into play through colors only. Both the ELG_LOP and ELG_VLO selections are selected with a $g_{\text{fib}} < 24.1$ cut, hence the two selections present very similar g -band magnitude distributions. However, the different locations of the selection boxes in the $g - r$ versus $r - z$ diagram imply different magnitude distributions in the r and z bands, with the ELG_LOP targets being 0.2 mag fainter than the ELG_VLO targets in the r band and 0.5 mag fainter than those in the z band. We thus expect the ELG_LOP selection to have a stronger dependency with the imaging z -band depth.

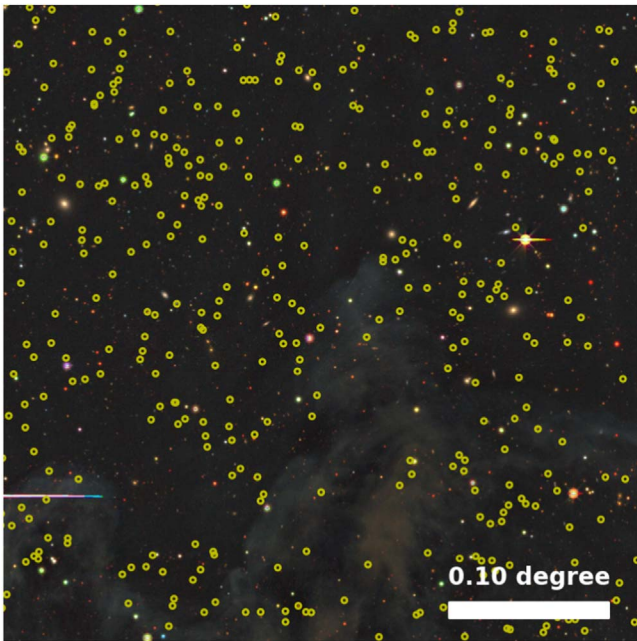


Figure 7. Example of a region with Galactic dust clouds causing extreme extinction variations at small scales, which are imprinted in the ELG_LOP target sample (yellow circles). The cutout is centered at (R.A., decl.) = (31°40, 20°66) and it is 0°4 wide.

5.2. Density Maps

Figure 6 displays the density fluctuations of the ELG_LOP (top) and ELG_VLO (bottom) targets across the whole LS-DR9 footprint. The 14,000 deg² footprint covered by DESI is indicated by the thick black contours. Several features are visible, especially for the ELG_LOP sample; we comment on the most noticeable ones.

The blue regions at (R.A., decl.) \sim (130°, 60°) or (R.A., decl.) \sim (30°, 20°) are underdensities due to a region of high Galactic dust, with many small-scale structures, as illustrated in Figure 7. A possible interpretation could be that even though the imaging is deeper in the dusty regions of the footprint, the extinction effect is only partially corrected in those regions (see Section 2.1)—in particular, great variations of the dust extinction at small scale cannot be handled at the imaging level. However, we note that some high-extinction regions can show an excess of ELG_LOP targets, as for instance at (R. A., decl.) \sim (345°, 20°). Proper explanations of these effects will likely require a detailed analysis of the interplay of the target selection with the dust extinction, the imaging depth, and the behavior of the *Tractor* source detection and fitting in those regions. Approaches like that of Obiwan (Kong et al. 2020), which injects fake sources into the imaging itself and then runs *Tractor*, may highlight the key information for such issues.

The ELG_LOP sample seems to have an overdensity along the Sagittarius Stream, displayed as a dashed line in Figure 6. The Sagittarius Stream has a stellar population that is bluer than the Galactic population, and it could in principle add contaminants to the ELG_LOP selection. As of now, it is not clear whether the ELG_LOP overdensity is due to the Sagittarius Stream stars or whether it is just concomitant: a detailed analysis of the spectroscopic observations will be required to clarify the issue.

The overdensities in the North at (R. A., decl.) \sim (180°, 40°) or (210°, 40°) correspond to regions of shallower extinction-corrected *g*-band imaging (see the bottom plot of Figure 1).

Last, we comment on three other features that are noticeable on these maps, even though they are well outside the DESI footprint and, hence, are not relevant for the DESI observations. For both selections, the density becomes slightly smaller below the decl. = -30° latitude in the DES region. This is due to a known shift of approximately 0.02 mag in the *z* band, where the calibration method transitions from Pan-STARRS 1 to UGSS (Padmanabhan 2008; D. J. Schlegel et al. 2023, in preparation). The large ELG_LOP overdensity at (R.A., decl.) \sim (80°, -60°), at the very south edge of the DES region, is contamination from the Large Magellanic Cloud, which adds a high density of blue stars in the region. And the ELG_LOP overdensity at (R.A., decl.) \sim (150°, -20°) is due to the much shallower imaging there (see Figure 1).

5.3. Photometric Systematics

In this section, we show how the ELG target selection depends on the imaging and foreground properties. As already stated, ELG targets have magnitudes close to the imaging depth, which makes the sample sensitive to fluctuations in the imaging and foreground maps.

We consider here the simplest set of maps. For the foreground, they encompass the Gaia stellar density and the Galactic dust extinction (the $E(B-V)$ parameter); besides, we also consider the projected distance to the Sagittarius Stream, as it could be a relevant quantity to consider, as seen in the previous section. For the imaging, in each of the three *grz* bands, we consider the seeing (the point-spread function or “PSF size”) and the “galaxy depth” corrected for dust extinction. We use dust extinction-corrected depths, as the target selection relies on dereddened magnitudes.

Figures 8 and 9 show how the ELG_LOP and ELG_VLO target selection densities vary with the foreground and imaging maps, for each of the North, South-DECaLS, and South-DES footprints. We note that we here discard the South-DES decl. $< -30^\circ$ region, because of the small photometric calibration issue mentioned in the previous section. These figures are based on 0.05 deg² HEALPix pixels ($n_{\text{side}} = 256$). For each footprint, the density variations are normalized to the average density over the footprint. The different properties of each footprint are clearly visible in this plot, and illustrate the need for them to be analyzed separately.

In general, both selections display similar trends, with the ELG_LOP sample showing stronger trends, as is expected from the fact that it contains fainter objects.

The most significant dependency involves the *g*-band depth, which shows two behaviors. For the North and the South-DECaLS footprints, the density decreases with increasing depth, whereas for the South-DES footprint, it increases with increasing depth. A possible explanation could be the following: for shallow imaging, the trend would be driven by contamination from stars and $z < 0.6$ galaxies, due to the scattering in the $g-r$ color, which makes them move inside the selection box. The scattering also affects $z > 0.6$ galaxies, making them move outside the selection box. However, the densities of such $z > 0.6$ galaxies are much smaller than those of the $z < 0.6$ galaxies (see Figure 3). The net effect would be an increase in the number of selected targets. The strength of this effect decreases as the depth increases (with the color

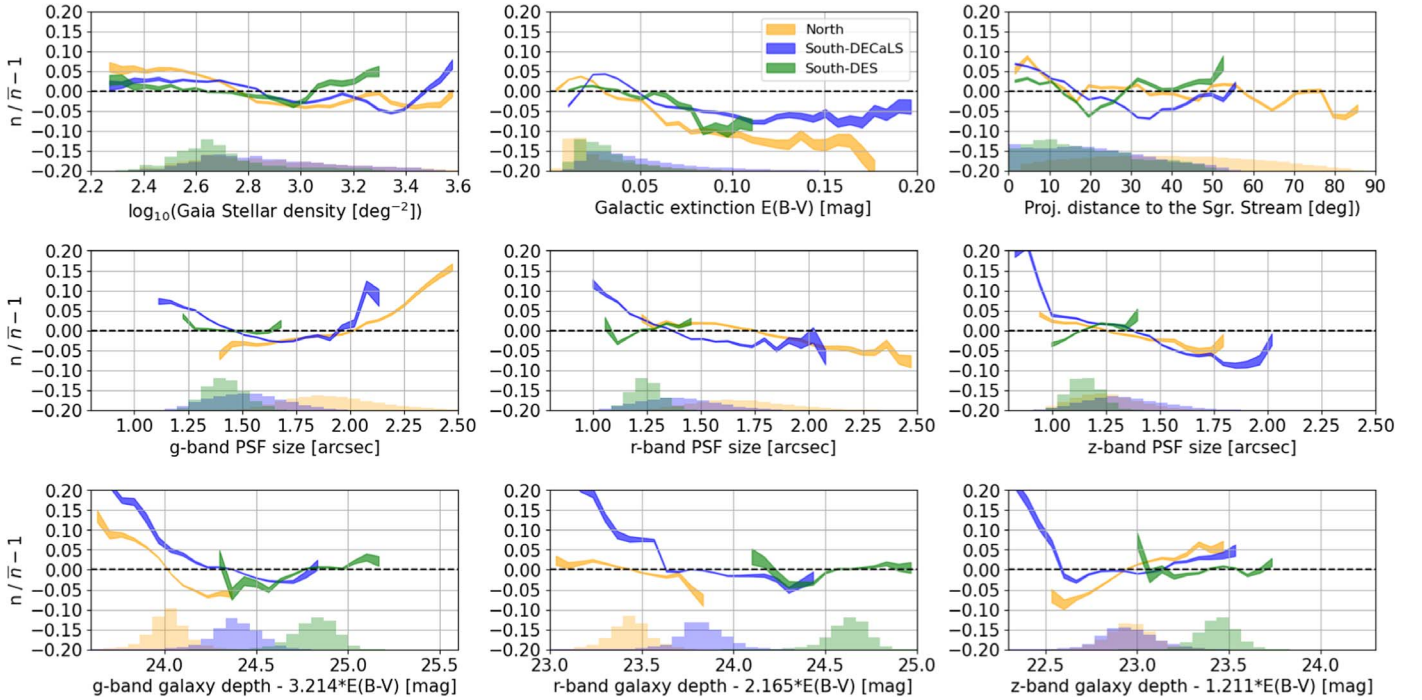


Figure 8. Main Survey ELG_LOP target density variations, with foreground (top) and imaging seeing (middle), as well as depth (bottom). The South-DES decl. $< -30^\circ$ region has been excluded. We consider bins with at least 100 HEALPix pixels.

scattering decreasing), and the impact on density eventually vanishes. Another effect could come into play with deeper imaging, namely the increase in the number of detected sources in the imaging: this second effect could explain the trend seen in the South-DES region.

There is a decrease in the target density with Galactic extinction for $E(B - V) > 0.05$ mag, this trend being stronger in the North than in the South-DECaLS and South-DES footprints. This could be explained by two facts: the imaging depth does not fully correct for the Galactic extinction (Section 2.1) or the regions with high extinction are embedded in small-scale structures that cannot be correctly accounted for in the imaging strategy (Figure 7).

Interestingly, both selections lead to increased density close to the Sagittarius Stream, which could be explained by contamination from stars from the stream.

5.4. Sensitivity to Photometric Zero-point Uncertainties

We estimate the sensitivity of the Main Survey ELG target selection to σ_{zp} , the imaging photometric zero-point uncertainties, using the same approach as in Myers et al. (2015) and Raichoor et al. (2017). The results are reported in Table 3.

In each of the g , r , and z bands, one at a time, we add ± 0.01 mag to the photometry and rerun the target selection algorithm to estimate $\delta N_{0.01} = \frac{|\Delta N|}{N}$, the fractional change in the target density due to this magnitude shift. We find consistent $\delta N_{0.01}$ values across the footprints. The ELG_LOP selection has $\delta N_{0.01} \sim 0.05$, 0.03, and 0.01 in the g , r , and z bands, respectively. The ELG_VLO selection has $\delta N_{0.01} \sim 0.04$, 0.05, and 0.04 in the g , r , and z bands, respectively. We notice that the selections have different sensitivities in the z band, ELG_VLO being more sensitive.

The expected rms variation in the number density due to shifts of the imaging zero-point is then estimated to be

$\frac{\delta N_{0.01}}{0.01} \times \sigma_{zp}$. LS-DR9 has σ_{zp} values of 0.003 mag in the g and r bands, and of 0.006 mag in the z band (D. J. Schlegel et al. 2023, in preparation). If we assume Gaussian errors for the zero-points, 95% of the footprint lies within $\pm 2\sigma_{zp}$ of the expected zero-point in any photometric band, meaning that 95% of the footprint has a variation in target density lower than $4 \times \sigma_{zp} \times \frac{\delta N_{0.01}}{0.01}$. The resulting fluctuations for each photometric band are given in Table 3. Both selections have density fluctuations of 1%–6% in all cases, except for the ELG_VLO sample in the z -band, where the density fluctuations are about 8%–9%. That level of fluctuation is reasonable, and should be able to be addressed with the weighting scheme in the LSS analysis.

6. Spectroscopic Data

We now present preliminary results from the DESI spectroscopic observations of this ELG sample. These observations include three phases of the DESI experiment: the SV, One-Percent, and Main Surveys. This section introduces these observations, along with the reduction of the data, which will be used in Section 7 to perform the analysis.

The SV and One-Percent data presented below will be part of the SV data that will be released in the DESI Early Data Release (DESI Collaboration et al. 2023b, in preparation).

6.1. The DESI Instrument

The DESI instrument, which is described in detail in DESI Collaboration et al. (2016b) and Abareshi et al. (2022), is a multi-spectroscopic instrument that is mounted at the prime focus of the 4 m Mayall Telescope at Kitt Peak, Arizona. The focal plane covers a field of view of about 8 deg^2 (T. Miller 2023, in preparation), and it is equipped with 5000 fiber positioners (Silber et al. 2022), distributed in ten “petals.” For each of the 500 fibers of a given petal, the light is dispersed by

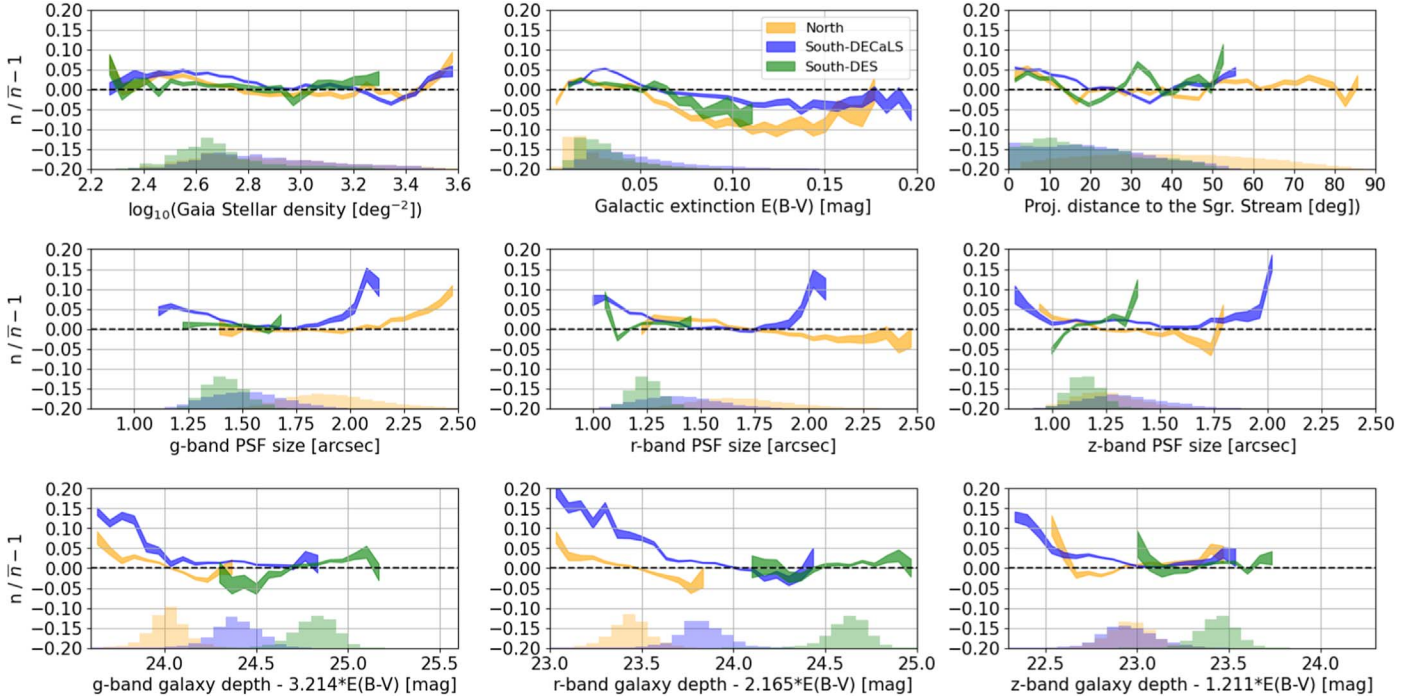


Figure 9. Main Survey ELG_VLO target density variations, with foreground (top) and imaging seeing (middle), as well as depth (bottom). The South-DES decl. $< -30^\circ$ region has been excluded. We consider bins with at least 100 HEALPix pixels.

one of the 10 three-arm spectrographs (“B”: 360 nm to 600 nm; “R”: 560 nm to 780 nm; and “Z”: 740 nm to 990 nm). The resolving power ($R = \lambda/\Delta\lambda$) increases with the wavelength, from ~ 2000 at the shortest wavelengths to nearly ~ 5500 at the longest ones. The wavelength coverage and the resolving power were designed to ensure that the instrument could measure and resolve the ELG [O II] $\lambda\lambda$ 3726,29 Å doublet in the $0.6 < z < 1.6$ range.

6.2. Observations

Here, we briefly summarize the DESI ELG spectroscopic observations that are used hereafter; the interested reader can find details in DESI Collaboration et al. (2023a, in preparation). DESI observations are conducted by “tiles,” i.e., by a group of 5000 fibered targets that are observed at once. Each tile is observed so as to reach a required average S/N value for all spectra. This is done through the computation of an effective exposure time, EFFTIME_SPEC , which accounts for the observing conditions and the per-fiber properties (Guy et al. 2022). The sky map of the DESI ELG tiles used in this paper is displayed in Figure 10, which shows that each program has a specific tiling coverage of the footprint.

The above data include 37 tiles with ELG targets from the SV (2020 December–2021 March), which explore extended samples in order for the target selection to be finalized (see Section 4). They typically have $\text{EFFTIME_SPEC} \sim 4000$ s—i.e., four times the nominal Main Survey EFFTIME_SPEC —with the result that the observations provide secure data for studying the faint end of the explored samples. Three of those tiles have much higher EFFTIME_SPEC values (7000–15,000 s) and were used to build a truth table of about 10,000 ELG spectra with VI (Lan et al. 2022).

The One-Percent Survey (2021 April–2021 May) observed 239 dark tiles distributed over 20 regions (“rosettes”) of the

NGC, with an EFFTIME_SPEC of about 1300 s, i.e., 30% larger than the nominal Main Survey EFFTIME_SPEC . A specificity of the One-Percent Survey observations is that most of the targets that did not have a conclusive z_{spec} after their first observation were reobserved with another tile, to increase the S/N. This significantly complicates the analysis in Section 7, as such repeat observations of the faintest targets, to secure a reliable z_{spec} measurement, are not representative of the Main Survey. In what follows, repeat observations—i.e., observations of the same target from different tiles—are thus removed from the One-Percent Survey analysis.

We use the Main Survey observations processed in Guy et al. (2022), which were taken from 2021 May to 2021 July and include 305 dark tiles with a narrow distribution of EFFTIME_SPEC (1100 ± 190 s). This data set, displayed in green in Figure 10, only covers parts of the North and South-DECaLS footprints. Last, for the redshift distribution (Figure 21) and the comparison with the HSC z_{phot} (Appendix B1), we complete this Main Survey sample with 973 Main Survey dark tiles, observed from 2021 September to 2021 December (in orange, in Figure 10), which provide significant coverage of the SGC, so that we have representative sampling of the three footprints (North, South-DECaLS, and South-DES), in particular in terms of the imaging depth, the Galactic extinction, and the stellar density. The pipeline reduction for this sample is not rigorously the same as that described in Section 6.5—it is slightly less advanced, but is still a very close version.

6.3. Main Survey Fiber Assignment

When designing a tile for observation, fibers are assigned to targets by means of a priority scheme. We provide here a brief presentation for the Main Survey, as this is relevant for the subsequent ELG analysis. The interested reader will find more

Table 3
Sensitivity of the Main Survey ELG Target Selection to the Imaging Photometric Zeropoint Uncertainties

Band	Footprint	σ_{zp} (mag)	ELG_LOP		ELG_VLO	
			$\delta N_{0.01}$	Fluctuations over 95% of the area (percent)	$\delta N_{0.01}$	Fluctuations over 95% of the area (percent)
g	North	0.003	0.052	6.2%	0.038	4.6%
	South-DECaLS		0.055	6.5%		4.6%
	South-DES		0.052	6.2%		4.8%
r	North	0.003	0.030	3.6%	0.046	5.5%
	South-DECaLS		0.036	4.3%		5.4%
	South-DES		0.032	3.8%		6.1%
z	North	0.006	0.004	1.0%	0.035	8.4%
	South-DECaLS		0.005	1.2%		7.7%
	South-DES		0.004	0.9%		9.0%

Note. Column (3): the imaging photometric zeropoint uncertainties (σ_{zp}). Columns (4) and (6): the fractional change in target density due to a ± 0.01 mag shift in the zeropoint ($\delta N_{0.01}$). Columns (5) and (7): the expected fluctuations in the number of selected targets over 95% of the footprint.

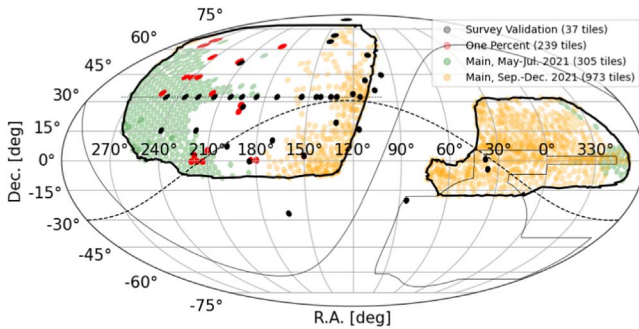


Figure 10. Sky distribution of the DESI-observed dark tiles used in this paper. They come from the following surveys: SV (black, 37 tiles), One-Percent (red, 239 tiles), and Main (2021 May–July: green, 305 tiles; 2021 September–December: orange, 973 tiles).

details in A. Raichoor et al. (2023, in preparation), for the fiber assignment process and statistics, and in E. F. Schlafly et al. (2023, in preparation), for the Main Survey tiling description and properties.

The Main Survey dark tiling consists of seven overlapping layers (“passes”), with each layer being made of nonoverlapping tiles. This tiling results in an average coverage of 5.2 tiles in the dark program.

In the dark-time observations, the ELG targets are at lower priority than the LRG and QSO targets, and thus they have a lower fiber assignment rate. The fiber assignment rates at the end of the survey are expected to be 0.99 for the QSO sample, 0.89 for the LRG sample, 0.69 for the ELG_LOP sample, and 0.42 for the ELG_VLO sample⁴³ (DESI Collaboration et al. 2023a, in preparation; A. Raichoor et al. 2023, in preparation). These values account for the 1% loss rate affecting the LRG and ELG observations, i.e., where observations are discarded because of nonvalid fibers (e.g., due to mechanical issues or petal rejection; see DESI Collaboration et al. 2023a, in preparation; E. F. Schlafly et al. 2023, in preparation). In the

⁴³ The ELG fiber assignment rates are lower than those for the other DESI tracers: we refer the interested reader to Bianchi et al. (2018), for methods of recovering unbiased correlation function measurements, despite such low fiber assignment rates.

event of such nonvalid fibers occurring for the QSO targets, we expect these targets to be reobserved in subsequent overlapping observations, as QSO targets have top priority.

In detail, for the ELG_LOP (ELG_VLO) sample, we expect fiber assignment rates of 0.69, 0.69, and 0.73 (0.43, 0.41, and 0.45) in the North, South-DECaLS, and South-DES footprints, respectively. The small variations are due to: (1) target density differences (see Table 1), with higher-density samples having a lower fiber assignment rate; and (2) slightly denser tiling in the South-DES footprint, because of the chosen tiling pattern (see E. F. Schlafly et al. 2023, in preparation).

6.4. ELG and QSO Targets

Unlike the intersection between the ELG and LRG target samples, which is virtually nil, the intersection between the ELG and QSO target samples is non-negligible. The fiber assignment scheme favors the QSO targets, and given that the “ELGxQSO” targets are not a representative subsample (they are brighter and have different colors), we need to account for this in the subsequent analysis. We stress that our goal here is only to correct for this overrepresentation of the ELGxQSO targets in the data set used in this paper, so that we can characterize the ELG sample in an unbiased way (in terms of its redshift success and redshift distribution). We do not perform any clustering analysis in this paper.

6.4.1. ELG and QSO Targets: Fiber Assignment

In the Main Survey, about one-third of the QSO targets are also ELG targets, mostly ELG_LOP ones. Those ~ 100 deg⁻² ELGxQSO targets represent about 5% of the ELG_LOP targets, and they will be overrepresented in the observations, getting more fibers in the first passes, and eventually being fully assigned, as they are also QSO targets. After each of the seven passes, the expected fiber assignment rates for the ELG_LOPxQSO sample are: 0.47, 0.74, 0.88, 0.95, 0.98, 0.99, and 1.00. For the overall ELG_LOP sample, they are expected to be: 0.07, 0.16, 0.27, 0.38, 0.49, 0.60, and 0.69, respectively. So, for instance, if our analysis uses an observed tile in a region with no previous observations, this ELG_LOPxQSO sample

will thus represent 35% of the observed ELG_LOP targets in this tile, and should be appropriately down-weighted.

6.4.2. ELG and QSO Targets: Weighting

In the analysis, one needs to account for the ELGxQSO targets being overrepresented in our observations. This is mostly critical for the SV data (excluding the 12 ELG-dedicated SV tiles) and the Main Survey data used in this paper (which mostly have one pass coverage; see Figure 10); the effect is less pronounced for the One-Percent observations, which have been completed and have a very high fiber assignment rate.

In the analysis in Section 7—except for the repeat analysis—we thus correct that effect, and appropriately down-weight the ELGxQSO observed targets. We split the sky into HEALPix pixels that are large enough⁴⁴ to reasonably track the density fluctuations of the selections. For each of the SV, One-Percent, and Main Survey selections, and each of the Main ELG_LOP, and ELG_VLO selections, we compute f_{targ} , the fraction of QSO targets in the considered selection for each pixel. If we note n_{QSO} and n_{notQSO} , the numbers of spectroscopically observed QSO and non-QSO targets for the considered selection in each pixel, we define the per-pixel weight to be applied to the n_{QSO} targets has follows: $w_{\text{QSO}} = f_{\text{targ}} \times n_{\text{notQSO}} / (n_{\text{QSO}} - f_{\text{targ}} \times n_{\text{QSO}})$. This ensures that $w_{\text{QSO}} \times n_{\text{QSO}} / (n_{\text{notQSO}} + w_{\text{QSO}} \times n_{\text{QSO}}) = f_{\text{targ}}$, i.e., that the weighted ELGxQSO observed targets represent the same fraction of the observed sample as the parent target sample, for each pixel.

6.5. Data Reduction

The spectroscopic data reduction and the redshift fitting with the Redrock software⁴⁵ are fully described in Guy et al. (2022) and S. J. Bailey et al. (2023, in preparation), respectively.

We only report on the sky subtraction performance, which is important for the ELG spectra, as those are at low S/N and the emission lines are used for redshift identification; this is especially relevant at $z > 1.5$, where the [O II] doublet falls in a region with many sky emission lines. Guy et al. (2022) demonstrate that the pipeline performs very well overall for typical dark-time exposures. For the continuum region, the rms of the spectral residuals in the sky fibers after sky subtraction is consistent to better than 1% with the expected noise from the continuum. For the brightest sky emission lines, it is at its worst, at the order of 3%; nevertheless, the model variance accounts for this effect, ensuring that the optimization for the redshift fitting is properly done.

We discard any observed spectrum with flagged issues in the data (COADD_FIBERSTATUS \neq 0). A key output quantifying the reliability of the best-fit z_{spec} is the χ^2 difference between the best-fit template and the second-best one (DELTACHI2). A large DELTACHI2 generally implies a reliable z_{spec} measurement.

We also use the S/N of the measured [O II] flux (FOII_SNR), which is computed as follows. The continuum is estimated in the vicinity of the doublet, from the wavelengths 200 Å (in rest frame) blueward of the [O II] doublet. The [O II] doublet is then simply fitted with two Gaussians, at the expected positions corresponding to the measured z_{spec} . The

[O II] flux, the line ratio, and the line width remain free in the fit.

Figure 11 compares the DESI spectrum of an ELG target to the one observed with the eBOSS survey. The typical eBOSS observations were for 1 hr, in contrast to the 15 minutes for a typical DESI observation. This is a representative $z_{\text{spec}} \sim 0.85$ ELG spectrum, with mostly undetected continuum and some significant emission lines. The zoom-in panels on the bottom row show the improvements resulting from DESI: its higher resolution allows one to nicely resolve the [O II] doublet, which provides an unambiguous feature for estimating the redshift, plus it also provides sharper emission lines, with higher S/Ns.

7. Spectroscopic Properties of the Main Sample

This section presents the spectroscopic properties of the ELG targets, based on the analysis of the spectroscopic data presented in Section 6.

7.1. Reliable z_{spec} Criterion

We introduce the criterion that is hereafter used to select a reliable z_{spec} for the ELG spectra, which is a cut in the {FOII_SNR, DELTACHI2} space. We emphasize that this is a simply preliminary criterion, to show at first order what can be achieved.

For each DESI tracer, the ELG spectra require a reliable dedicated z_{spec} measurement criterion, which maximizes the fraction of selected redshifts and minimizes the fraction of catastrophic redshifts in the selected sample, typically at the 1% level. Such requirements are driven by the LSS analysis, which is sensitive to catastrophic redshifts. One possible criterion is a high DELTACHI2 value, as used for other DESI tracers (Chaussidon et al. 2022; Hahn et al. 2022; Zhou et al. 2022). The specificity of the ELG spectra results from their being at low S/N: a z_{spec} that is reliably estimated with the [O II] doublet may not have a large DELTACHI2 value, as the pixels relating to the [O II] doublet represent a marginal fraction of the pixels, and a fit with a single emission line with a different redshift could still provide a comparable χ^2 . For this reason, selecting a reliable z_{spec} with a DELTACHI2 criterion would only discard a large fraction of good redshifts, noticeably at high redshift, where the [O II] doublet is the only feature in the spectrum.

A relevant parameter space to consider in this regard is the {FOII_SNR, DELTACHI2} space. Figure 12 shows the criterion that we adopt in this paper:

$$\log_{10}(\text{FOII_SNR}) > 0.9 - 0.2 \times \log_{10}(\text{DELTACHI2}). \quad (1)$$

This figure is computed using approximately 3500 ELG_LOP and ELG_VLO VI spectra. For these spectra, the VI provides two pieces of information from the deep reductions: $z_{\text{spec,VI}}$ and QA_{VI} , its confidence level—both these quantities are merged from the diagnosis of several inspectors. The VI confidence level QA_{VI} ranges from 0 to 4, and, following the definition established in Lan et al. (2022), spectra with $\text{QA}_{\text{VI}} \geq 2.5$ are considered to provide a robust $z_{\text{spec,VI}}$, which we consider as the truth here. We conservatively consider all spectra with $\text{QA}_{\text{VI}} < 2.5$ to be failures in shallower reductions. As these observations come from the three deep SV VI tiles, which have been exposed to many exposures, we are able to generate several tens of coadded reductions with EFFTIME_SPEC ranging from 200 to 1600 s. Then, for each z_{spec} measurement

⁴⁴ $n_{\text{side}} = 16$, i.e., a pixel area of 13.4 deg².

⁴⁵ <https://github.com/deshub/redrock>

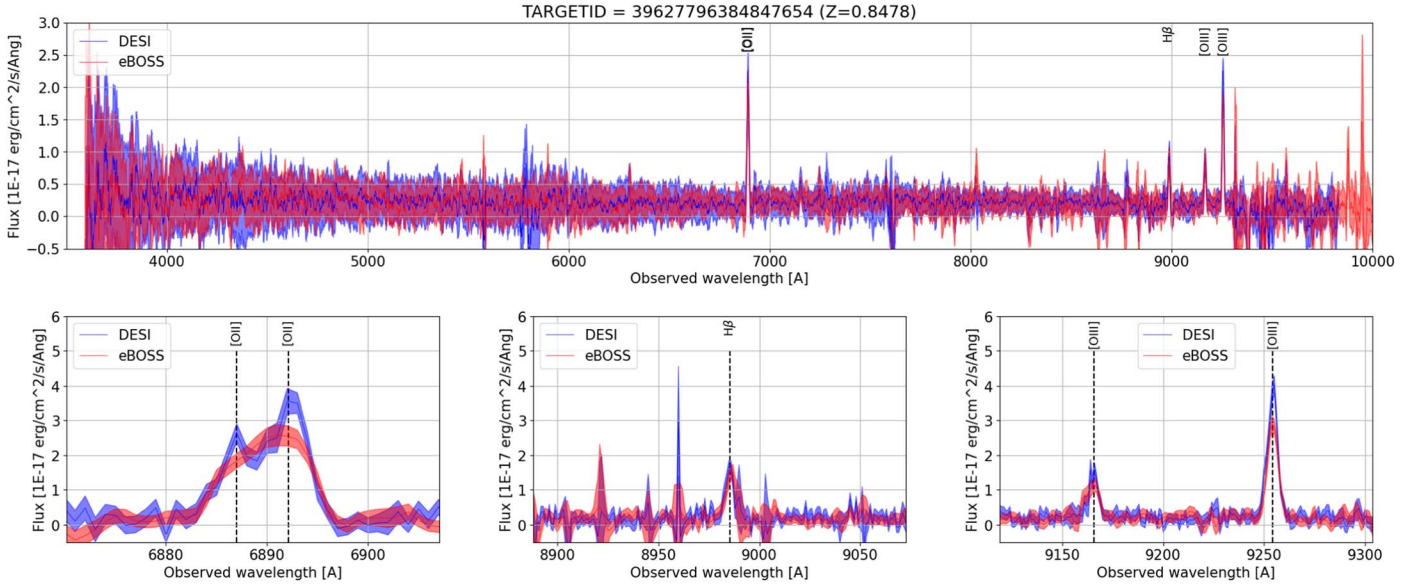


Figure 11. Typical ELG spectrum, observed with eBOSS over ~ 1 hr (red) and DESI over ~ 15 minutes (blue). The ELG target has $g_{\text{fib}} = 23.2$ mag and $z_{\text{spec}} = 0.8478$. The top panel shows the full observed wavelength range, while the bottom three panels show zoom-ins on the main emission lines for the spectrum: the [O II] doublet (left), the H β line (middle), and the [O III] $\lambda\lambda$ 4960, 5008 Å lines (right). The spectra in the top panel are smoothed with a box of 11 pixels. The shaded regions show the estimated 1σ uncertainty on the measured flux.

from those shallower reductions, we compare its value to the $z_{\text{spec,VI}}$ from the deep reductions. We consider the z_{spec} measurement to be “VI-validated” if it verifies the two following criteria:

$$QA_{\text{VI}} \geq 2.5, \quad (2a)$$

$$c \cdot |z_{\text{spec}} - z_{\text{spec,VI}}| / (1 + z_{\text{spec,VI}}) < 1000 \text{ km s}^{-1}, \quad (2b)$$

where c is the speed of light in km s^{-1} . The top panel of Figure 12 shows that a simple cut in DELTACHI2 is not at all optimal: a sample selected with the fiducial DELTACHI2 > 9 threshold (used in Redrock to flag a low-reliability z_{spec} measurement) would be highly contaminated with catastrophic z_{spec} measurements; a more conservative threshold, e.g., DELTACHI2 > 25 , would discard a significant number of reliable z_{spec} measurements. Our simple criterion of Equation (1) selects more than 95% of the reliable z_{spec} measurements, while keeping a very low fraction of catastrophic z_{spec} (about 1%).

The bottom panel of Figure 12 displays the average redshift for each position in the $\{\text{FOII_SNR}, \text{DELTACHI2}\}$ plane, using the $z_{\text{spec,VI}}$ measurements with $QA_{\text{VI}} \geq 2.5$. It shows that the high-redshift ELG targets, which are the most valuable ones for DESI, have low DELTACHI2 values, despite their reliable z_{spec} measurement (likely from the identification of the resolved [O II] doublet).

A detailed analysis of all the recent Main Survey spectra—possibly enhanced with additional VI—will allow the refinement of this Equation (1) criterion. The likely improvements would be refining the cut in the space to enlarge the fraction of selected z_{spec} ; refining the selection in the $1.5 < z < 1.6$ range, where the [O II] doublet falls in the forest of sky emission lines; or refining it at $z < 1$, using other lines, like [O III] $\lambda\lambda$ 4960, 5008 Å.

7.2. FOII_SNR and z_{spec}

In this section, we discuss how the FOII_SNR and $S/N = 7$ [O II] flux limits vary with redshift, with a focus on the $z > 1.5$ redshift range.

In Figure 13, we illustrate the how z_{spec} measurements selected by Equation (1) are distributed in the $\{\text{FOII_SNR}, Z\}$ plane, for the $0.6 < z < 1.6$ range (top), and by zooming in the $1.45 < z < 1.55$ range (bottom). To obtain the largest sample size, this figure displays about 600,000 Main ELG spectra observed in the SV and Main Survey from coadded reductions with $800 \text{ s} < \text{EFFTIME_SPEC} < 1200 \text{ s}$.

A noticeable feature is the drop in the measured FOII_SNR for some z_{spec} values, especially at $z > 1.5$, which corresponds to the [O II] doublet moving redward of 9300 Å. This is a consequence of the combination of two effects. On the one hand, the DESI throughput drops at wavelengths redder than 9300 Å, because of a lower quantum efficiency, together with more extinction in the atmosphere (see Figures 18 and 27 of Abareshi et al. 2022). On the other hand, the sky is densely populated by bright emission lines at wavelengths redder than 9300 Å, which leaves sky emission lines subtraction residuals in the ELG spectra, increasing the noise. We estimate that the two effects approximately have the same contribution to the drop in FOII_SNR seen at $z > 1.5$.⁴⁶

We illustrate the contribution of sky emission lines in two ways. First, we display in Figure 13 the sky emission line

⁴⁶ For a set of 26 EFFTIME_SPEC = 1000 s exposures, we consider these two quantities: $T/\sqrt{\text{sky}}$ for the throughput (i.e., the throughput divided by the square root of the measured sky) and the estimated noise in the sky-subtracted flux in the sky fibers for the sky emission line subtraction residuals. For each quantity, we compute the ratio of the value at $8900 \text{ \AA} < \lambda < 9300 \text{ \AA}$ to the value at $9300 \text{ \AA} < \lambda < 9700 \text{ \AA}$, for 100 realizations, where for each realization we take the mean value over consecutive 20 Å in a random position in the wavelength window, thus obtaining a representative value for the [O II] doublet at $z > 1.3$. We find a ratio of 0.60 for the throughput quantity and a ratio of 0.63 for the residual quantity. This amounts to a total drop in $\log_{10}(\text{FOII_SNR})$ of $\log_{10}(0.60 \times 0.63) = -0.42$, in broad agreement with the drop in Figure 13.

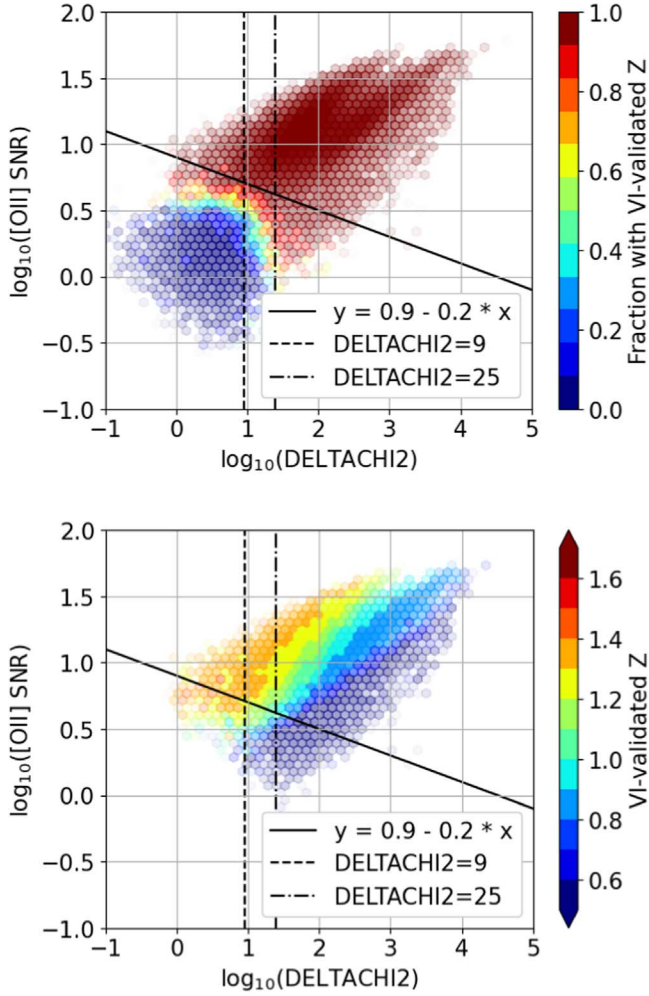


Figure 12. ELG target properties in the $\{\text{FOII_SNR}, \text{DELTACHI2}\}$ plane. Top: the fraction of z_{spec} validated by VI. Bottom: the average $z_{\text{spec,VI}}$. In both plots, the slanted solid line is our criterion for selecting reliable z_{spec} measurements, while the dashed and dotted-dashed vertical lines illustrate two threshold values for a lower cut in DELTACHI2 . The symbol transparency scales with the logarithm of the density. The data come from several reductions with $200 \text{ s} < \text{EFFTIME_SPEC} < 1600 \text{ s}$ of the 3500 ELG_LOP and ELG_VLO targets with VI spectra (see the text).

wavelengths converted into redshifts at the tops of the plots, assuming the redshift to be that of an [O II] doublet appearing at the same observed wavelength as the sky emission line, i.e., $z_{\text{spec}} = \lambda_{\text{sky}}/3728 - 1$. We then display in Figure 14 typical [O II] doublet measurements for ELG_LOP spectra: the left and middle columns each show four spectra at redshifts 1.3 and 1.55, respectively, with typical FOII_SNR values for the considered redshifts. One can clearly see the increase of the flux uncertainty at the sky emission line positions (the vertical blue dotted lines), with the effect of reducing the FOII_SNR . However, the [O II] doublet can still be successfully identified in the $z > 1.5$ region, thanks to the quality of the data and the pipeline. The right column of Figure 14 displays the spectra for the same four $z = 1.55$ targets, but with the deep observations available in the SV, nicely illustrating the unambiguous [O II] doublet. Future data reduction improvements (sky subtraction and [O II] flux measurement) will be valuable for

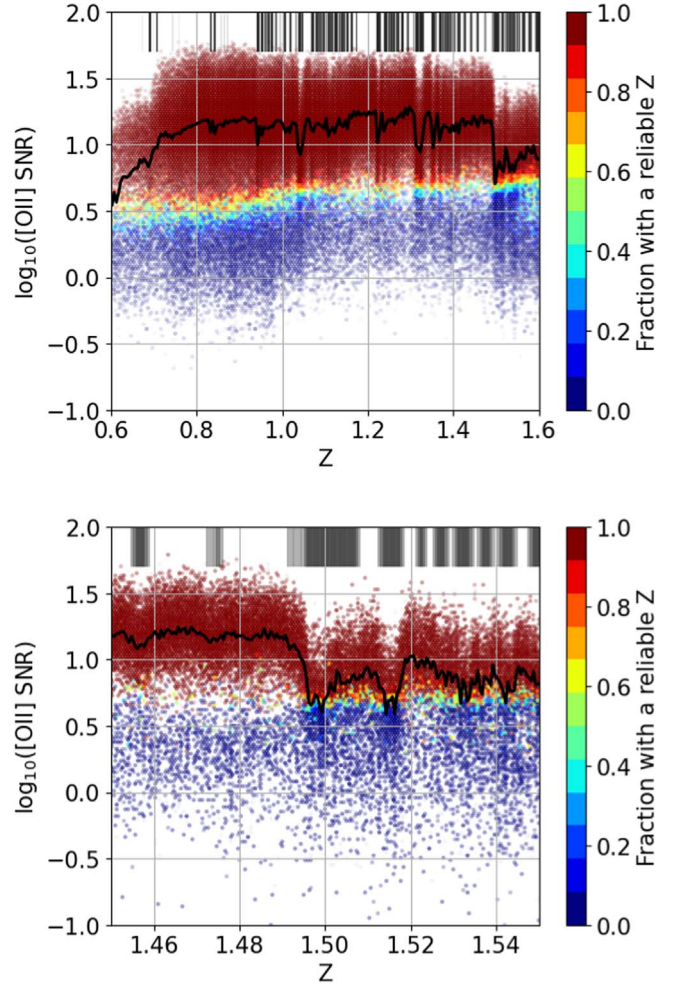


Figure 13. Distribution of the fraction of reliable z_{spec} measurements in the $\{\text{FOII_SNR}, Z\}$ plane, for the $0.6 < z < 1.6$ range (top), zooming in on the $1.45 < z < 1.55$ range (bottom). The solid black line displays the median FOII_SNR value. The symbol transparency scales with the logarithm of the density. In both plots, the sky emission lines are displayed as black rectangles at the top of the plots, where the sky wavelength λ_{sky} is converted to $z_{\text{spec}} = \lambda_{\text{sky}}/3728 - 1$ (that is, to the redshift of an [O II] doublet observed at that wavelength). The data come from about 600,000 Main ELG spectra observed in coadded reductions with $800 \text{ s} < \text{EFFTIME_SPEC} < 1200 \text{ s}$.

the analysis of the ELG sample, hopefully increasing the fraction of reliable z_{spec} at $z > 1.5$.

An interesting related measurement is the median $S/N = 7$ [O II] flux limit as a function of redshift. Indeed, a DESI requirement that played an important role in the instrumental design was that this limit should be 10, 9, 9, 8, and $9 \times 10^{-17} \text{ erg s}^{-1} \text{ cm}^{-2}$ in the redshift bins 0.6–0.8, 0.8–1.0, 1.0–1.2, 1.2–1.4, and 1.4–1.6, for nominal ELG observations (see Table 5 of Abareshi et al. 2022). We thus isolate the subsample with $6.8 < \text{FOII_SNR} < 7.2$ (this selection virtually fully passes our Equation (1) criterion), and display in Figure 15 its median [O II] flux as a function of redshift. The patterns relating to sky emission lines and throughput are naturally also visible here. The median values from these data are 7.1, 6.4, 7.0, 6.5, and $7.8 \times 10^{-17} \text{ erg s}^{-1} \text{ cm}^{-2}$ in the 0.6–0.8, 0.8–1.0, 1.0–1.2, 1.2–1.4, and 1.4–1.6 redshift bins. These values are on average 20% better than the requirements, illustrating the good performance of the DESI instrument.

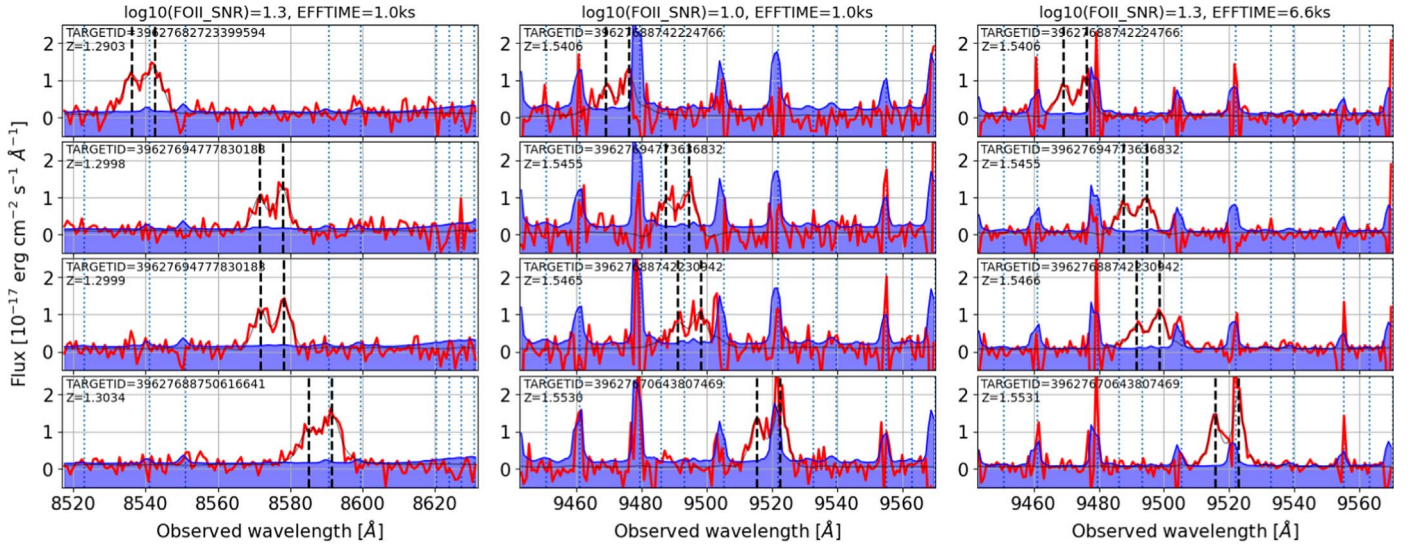


Figure 14. Illustration of the [O II] doublet in the ELG_LOP spectra. The left (middle) panel displays four spectra at $z = 1.3$ ($z = 1.55$) for $\text{EFFTIME_SPEC} = 1.0$ ks, with typical FOII_SNR values for the considered redshift. The right panel displays the same four $z = 1.55$ targets, but as observed with $\text{EFFTIME_SPEC} = 6.6$ ks. The thick red line shows the observed flux, the blue-filled region shows the flux uncertainty, and the black line is the best-fit model. The [O II] doublet wavelengths are displayed as vertical dashed black lines. The sky emission line wavelengths are indicated with the vertical blue dotted lines.

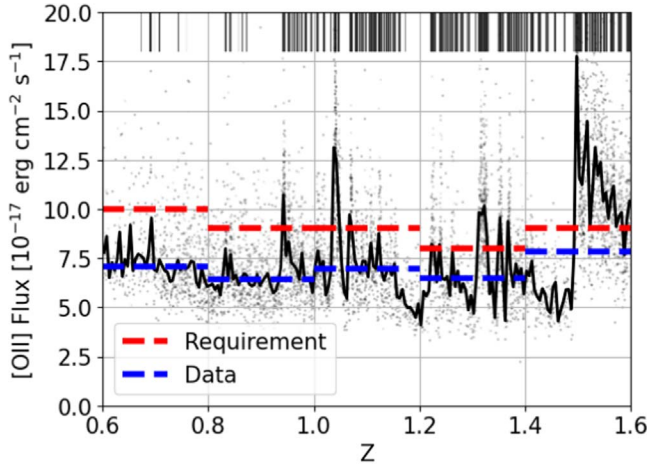


Figure 15. [O II] flux limit as a function of z_{spec} for $\text{FOII_SNR} = 7$ ELG spectra. Each point marks an ELG spectrum, with the solid black line displaying the median values. The horizontal dashed blue lines are the median values for the $dz = 0.2$ bins, while the horizontal dashed red lines are the DESI requirements. The sky emission lines are displayed as the black rectangles at the tops of the plots, as in Figure 13. The data come from 15,400 Main ELG spectra, observed in coadded reductions with $800 \text{ s} < \text{EFFTIME_SPEC} < 1200 \text{ s}$ and $6.8 < \text{FOII_SNR} < 7.2$.

7.3. Fraction of Selected Catastrophic z_{spec} Estimated from VI

An important quantity to control is the fraction of z_{spec} selected with our reliability criterion (Equation (1)), which has a catastrophic z_{spec} estimate.

We first assess this fraction using the SV VI sample from the three deep ELG tiles. We identify as a catastrophic z_{spec} estimate any measurement that passes our Equation (1) criterion, but fails Equation (2a) or Equation (2b). Here, we restrict the sample to spectra that would be selected for an LSS analysis, i.e., redshifts in the $0.6 < z < 1.6$ range passing our criterion (about 2900 ELG_LOP targets and 700 ELG_VLO targets). From the multiple reductions, we have at hand about 60 reductions spanning EFFTIME_SPEC values between 200 and 1600 s.

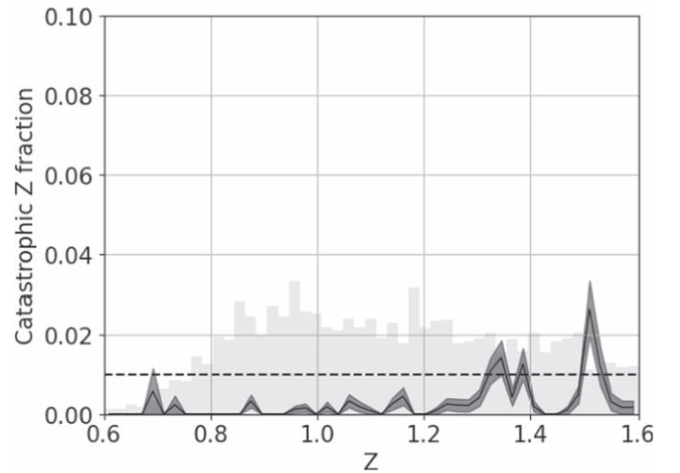
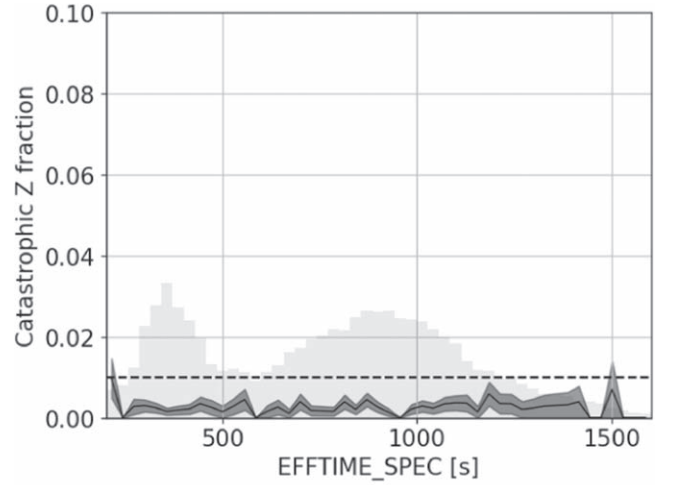


Figure 16. Fraction of catastrophic z_{spec} measurements for the Main ELG_LOP targets with $0.6 < z < 1.6$ and a reliable z_{spec} (Equation (1)), as a function of EFFTIME_SPEC (top) and z_{spec} (bottom). The gray histograms show the distributions of the probed values, and the dashed lines highlight a catastrophic z_{spec} fraction of 1%. The data come from different reductions of the 2900 ELG_LOP targets with VI.

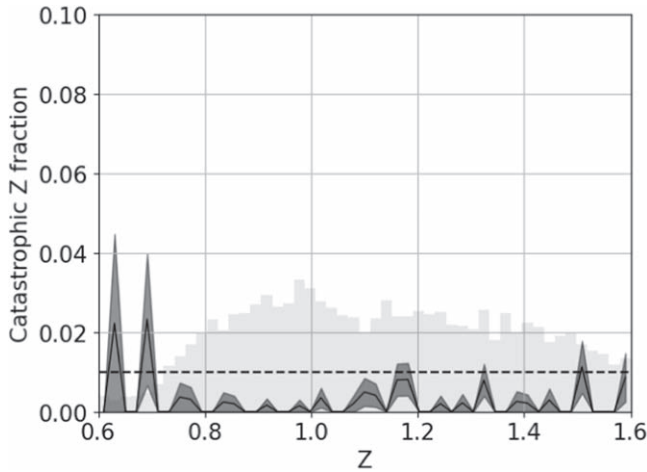


Figure 17. The same as Figure 16, but with the data coming from independent repeat observations from the SV, One-Percent, and Main Surveys (see the text for more details).

Figure 16 presents, for the ELG_LOP sample, the fraction of catastrophic z_{spec} as a function of EFFTIME_SPEC (top) and z_{spec} (bottom) values, and demonstrates that our criterion is effective in keeping the catastrophic z_{spec} fraction at the order of the percent level. For a typical $\text{EFFTIME_SPEC} \sim 1000$ s—the nominal EFFTIME_SPEC for the Main Survey—the catastrophic z_{spec} fraction for the ELG_LOP sample is $\sim 0.2\%$; for the ELG_VLO sample, it is virtually zero. The fraction is independent of EFFTIME_SPEC ; this is the desired behavior—i.e., a shallow reduction would naturally select many fewer spectra, but they would be of similar quality as spectra from a deeper reduction. As a function of z_{spec} , the catastrophic z_{spec} fraction is very low for $z < 1.2$, but it starts to increase for $1.3 < z < 1.4$, and is more significant for $1.5 < z < 1.6$.

The reasons are twofold. First, this reflects the fact that as the redshift increases, the emission lines move redward, leaving only the [O II] doublet. Then, for $1.5 < z < 1.6$, the [O II] doublet falls in a region with many strong sky emission line subtraction residuals (see Figure 13), likely preventing the VIs from securely confirming the redshift; our conservative choice of considering all the redshifts from a target with $\text{QA}_{\text{VI}} < 2.5$ as a failure (see Equation (2a)) likely explains the high values of the catastrophic z_{spec} fraction in $1.5 < z < 1.6$ (see the next section). In any case, to reduce the catastrophic z_{spec} fraction in the $1.5 < z < 1.6$ range, it will be necessary to improve the sky subtraction in the reduction pipeline.

7.4. Fraction of Selected Catastrophic z_{spec} and z_{spec} Precision Estimated from Repeats

We use repeat observations of the Main ELG targets to reassess the catastrophic z_{spec} fraction with an independent method, and to determine the z_{spec} measurement precision.

We build a sample of about 19,000 (5000) pairs of independent repeat observations of the Main ELG_LOP (ELG_VLO) targets, as follows. We consider the SV, One-Percent, and Main per-night reductions with $800 \text{ s} < \text{EFFTIME_SPEC} < 1200 \text{ s}$, each reduction being made from independent observations (either different tile observations or different exposures for a given tile). We then restrict the sample to the Main ELG targets with a reliable z_{spec} measurement (Equation (1)), and with one of the two

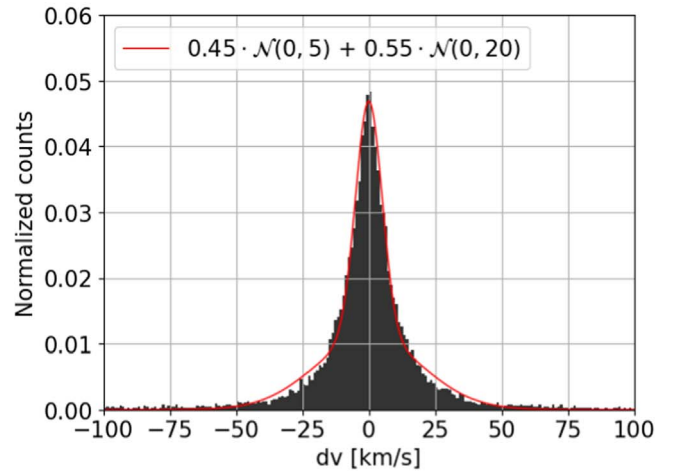
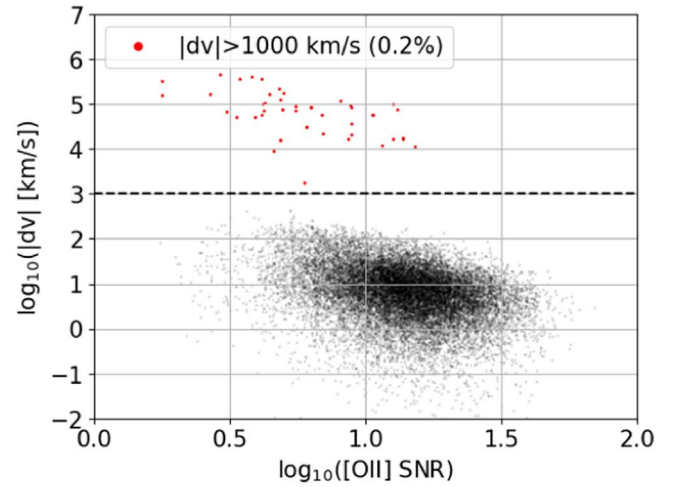


Figure 18. Redshift difference dv for 19,000 pairs of independent repeat observations of Main ELG_LOP targets with $0.6 < z < 1.6$. Top: dv as a function of FOII_SNR ; the black dashed line ($dv = 1000 \text{ km s}^{-1}$) illustrates the threshold used to identify the pairs with catastrophic measurements (red dots). Bottom: normalized distribution of the dv values; the red curve illustrates a reasonable representation of the distribution, from the weighted sum of two Gaussian distributions, with widths of 5 and 20 km s^{-1} . The data come from independent repeat observations from the SV, One-Percent, and Main Surveys (see the text for more details).

measurements in $0.6 < z < 1.6$, before identifying targets having two or more reductions. For each pair, we consider the redshift difference $dv = c \cdot (z_0 - z_1)/(1 + z_0)$, where c is the speed of light in km s^{-1} .

First, Figure 17 independently reassesses the catastrophic z_{spec} fraction as a function of redshift, using these repeat observations. We obtain subpercent fractions for all redshifts. This is consistent with Figure 16, except for the $1.3 < z < 1.6$ range, where the fraction from the repeats has much less pronounced peaks. This is consistent with our statement in previous section, that the peaks in Figure 16 are likely driven by our conservative choice of considering all redshifts from a target with $\text{QA}_{\text{VI}} < 2.5$ as failures.

The top panel of Figure 18 shows dv as a function of FOII_SNR for the ELG_LOP sample: only 0.2% of the pairs have a catastrophic measurement (red dots). The ELG_VLO sample has virtually zero catastrophic measurements. This fraction is in agreement with the catastrophic rate estimated in Section 7.3, from a totally independent method.

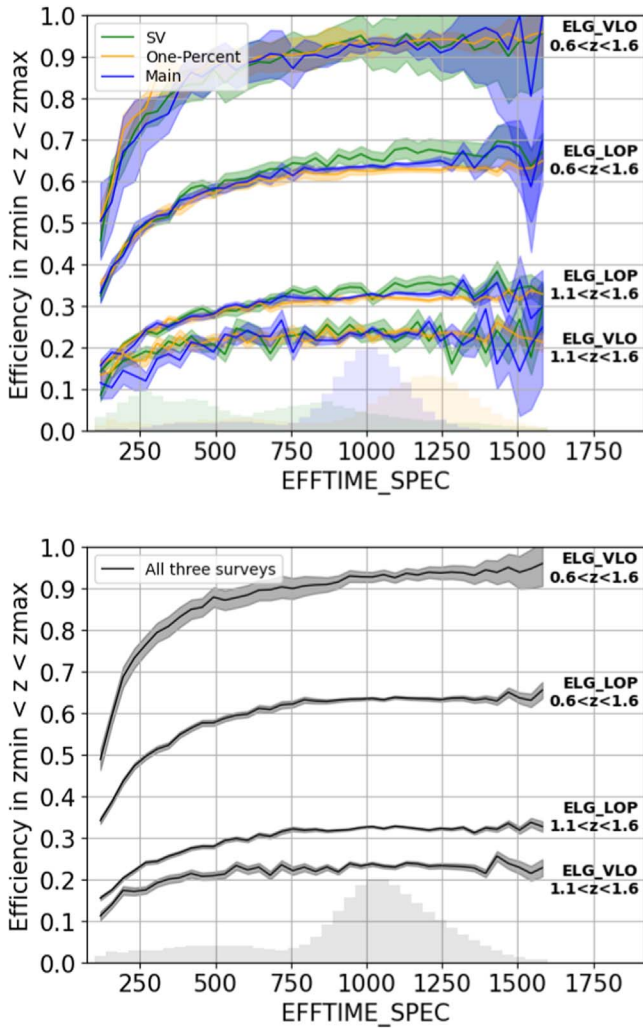


Figure 19. ELG Main selection efficiency in the $0.6 < z < 1.6$ and $1.1 < z < 1.6$ ranges, as a function of EFFTIME_SPEC , for the three surveys (the SV, One-Percent, and Main Surveys). Top: split per survey. Bottom: all surveys together. The histograms show the EFFTIME_SPEC normalized distributions.

Besides, this panel emphasizes that the $[\text{O II}]$ doublet is crucial for the ELG z_{spec} measurement, as it shows a clear correlation between the redshift precision and the FOII_SNR . The dv distribution is reported in the bottom panel of Figure 18, and it can reasonably be modeled by the weighted sum of two Gaussian distributions centered on zero with widths of 5 and 20 km s^{-1} . Eventually, if we use the same statistical measurement as Lan et al. (2022), we measure for the ELG_LOP (ELG_VLO) sample, $\text{MAD}(dv) \times 1.48/\sqrt{2} \sim 7 \text{ km s}^{-1}$ (9 km s^{-1}), where MAD is the median absolute deviation, in agreement with Lan et al. (2022). This z_{spec} precision is the best among the DESI tracers (Alexander et al. 2022; Lan et al. 2022), as the z_{spec} is based on sharp emission lines.

7.5. Efficiency: Fraction of Selected $z_{\min} < z_{\text{spec}} < z_{\max}$

We present the fraction of ELG spectra selected by our reliability criterion from Equation (1). Figure 19 displays that fraction as a function of EFFTIME_SPEC , for the Main ELG_LOP and ELG_VLO samples, and for the three surveys (SV, One-Percent, and Main; top panel). Combining those three surveys (bottom panel) allows us to probe a wide range of

EFFTIME_SPEC , with SV exploring the entire range, but with low statistics, and One-Percent probing values from 1000 to 1500 s, right in the high tail of the Main Survey range. For this figure, we restrict the sample to reductions with $100 \text{ s} < \text{EFFTIME_SPEC} < 1600 \text{ s}$. As demonstrated in Section 5.3, the ELG_LOP and ELG_VLO selections have some dependencies with the g -band imaging depth, the Galactic extinction, and the distance to the Sagittarius Stream. To allow a comparison of the three surveys, we: (1) restrict the sample to regions with g -band depth less than 24.5 mag and $E(B-V)$ less than 0.1 mag; and (2) subsample the SV and One-Percent data, so that the distances to the Sagittarius Stream values are representative of the distribution probed by the Main Survey.

We call the efficiency for a given $z_{\min} < z < z_{\max}$ range the fraction of observed ELG targets that obtain a reliable z_{spec} measurement in that redshift range. In this figure, we display the efficiency for the two important redshift ranges to control for the target sample—namely, $0.6 < z < 1.6$, the nominal redshift range, and $1.1 < z < 1.6$, the high-redshift part of that range, where ELGs are the most important tracers for DESI. For the Main Survey nominal $\text{EFFTIME_SPEC} \sim 1000 \text{ s}$, the ELG_LOP selection has an efficiency in $0.6 < z < 1.6$ of 60%–65%, and an efficiency in $1.1 < z < 1.6$ of 30%–35%. The ELG_VLO selection has a much higher efficiency in $0.6 < z < 1.6$, of 90%–95%, but an efficiency of only 20%–25% in $1.1 < z < 1.6$, as expected from the designed photometric cuts.

A noticeable feature in Figure 19 is the overall agreement between the three surveys. This highlights the relevance of DESI’s approach, using the SV to explore a large selection sample, then using One-Percent to refine the selections and check them with observations that are slightly deeper than nominal, before finalizing the Main Survey selection. And this is what allows us to combine the three surveys together (the bottom panel of Figure 19), leading to precise measurements over a large range of EFFTIME_SPEC values.

The second visible feature is the flattening of the efficiency curves toward large EFFTIME_SPEC values. As expected, the efficiency is low for low- EFFTIME_SPEC values, as the ELG spectra do not have high enough S/Ns to securely measure a z_{spec} value for most of the sample. The efficiency then strongly increases with increasing EFFTIME_SPEC , up to approximately 750 s. Finally, the efficiency flattens for EFFTIME_SPEC values larger than 750 s. As a consequence, the efficiency is rather constant over the EFFTIME_SPEC range probed by the Main Survey (the blue histogram in the top panel), which naturally has some scatter around the requested value of 1000 s.

7.6. Efficiency in the Photometric Space

This section analyzes the redshift efficiency in the $0.6 < z < 1.6$ and $1.1 < z < 1.6$ ranges as a function of the g -band magnitude and the grz -band colors. We use the SV ELG sample in order to quantify how the efficiency varies in the photometric space for the Main ELG selection, and at the borders of that selection. Such a study is used to finalize the Main Survey ELG selection.

We use reductions of the 25 SV ELG-only tiles with $800 \text{ s} < \text{EFFTIME_SPEC} < 1200 \text{ s}$, i.e., EFFTIME_SPEC values that are representative of the Main Survey. These tiles include about 43,200 ELG targets. The results are presented in Figure 20, where the top (bottom) row is the efficiency in the

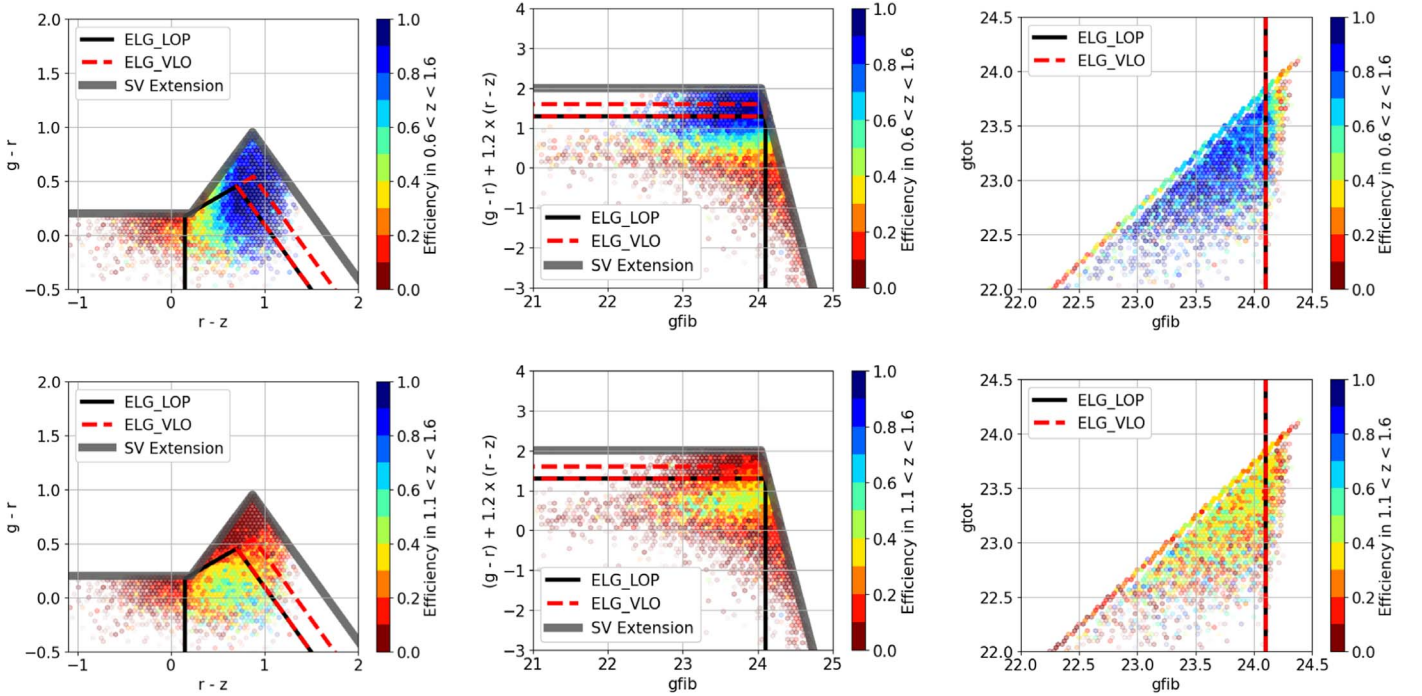


Figure 20. Efficiency as a function of color and magnitude, in the $0.6 < z < 1.6$ (top) and $1.1 < z < 1.6$ (bottom) range. The left column shows the $g-r$ vs. $r-z$ diagram, the middle column shows the $(g-r) + 1.2 \times (r-z)$ vs. g_{fib} diagram, and the right column shows the g_{tot} vs. g_{fib} plane. The Main ELG_LOP and ELG_VLO selections are displayed as the solid dark lines and the dashed red lines, respectively; the SV ELG selection is displayed as a thick solid gray line. The data come from different reductions of SV ELG-only tiles, with $800 \text{ s} < \text{EFFTIME_SPEC} < 1200 \text{ s}$, representative of the Main Survey EFFTIME_SPEC values.

Table 4
Efficiency per Sample and Footprint for the Main Survey ELG Selection

Sample	Footprint	Efficiency			
		All Redshifts	$0 < z < 0.6$	$0.6 < z < 1.1$	$1.1 < z < 1.6$
ELG_LOP	North	0.71	0.05	0.33	0.32
	South-DECaLS	0.68	0.05	0.29	0.34
	South-DES	0.73	0.02	0.31	0.39
ELG_VLO	North	0.93	0.01	0.70	0.21
	South-DECaLS	0.94	0.01	0.67	0.25
	South-DES	0.95	0.00	0.69	0.26

Note. We call the efficiency for a given $z_{\text{min}} < z < z_{\text{max}}$ range the fraction of observed ELG targets that provide a reliable z_{spec} measurement (according to Equation (1)) in that redshift range.

$0.6 < z < 1.6$ ($1.1 < z < 1.6$) range. The left column shows the $g-r$ versus $r-z$ diagram, the middle column shows the $(g-r) + 1.2 \times (r-z)$ versus g_{fib} diagram, and the right column shows the g_{tot} versus g_{fib} plane. On all plots, the Main ELG_LOP selection is displayed as a solid dark line, the Main ELG_VLO selection as a dashed red line, and the SV ELG selection as a thick solid gray line.

The $g-r$ versus $r-z$ plots include a $g_{\text{fib}} < 24.1$ selection, i.e., the same magnitude limit as the Main ELG sample. The efficiency in the $1.1 < z < 1.6$ interval (bottom panels) was the key motivation behind the Main ELG selection box definition. The efficiency inside the ELG_LOP selection box is rather stable, but sharply drops on the blue $r-z$ side (likely because any $z > 1.6$ galaxy cannot provide a reliable z_{spec}) and on the red $g-r$ side (likely due to contamination from stars and $z < 0.6$ galaxies). On the red $r-z$ side, the efficiency drops more smoothly, which justifies the ELG_VLO selection box definition, in conjunction with the very high efficiency of the

$g-r$ versus $r-z$ region for the $0.6 < z < 1.6$ range (top panels).

The $(g-r) + 1.2 \times (r-z)$ versus g_{fib} diagram allows us to test the effects of selecting ELG targets fainter than $g_{\text{fib}} < 24.1$. At fixed $(g-r) + 1.2 \times (r-z)$, the efficiency is actually rather stable for $(g-r) + 1.2 \times (r-z) > 0$, when going to magnitudes fainter than the Main ELG cut (the solid black line), illustrating the good performance of DESI. Nevertheless, such a selection would also bring in many failures from the $(g-r) + 1.2 \times (r-z) < 0.5$ region, which would mitigate the gain in efficiency; this, combined with the fact that going fainter increases the density variations with imaging and foreground properties (see Section 5.3), explains why it was not considered in the end.

The g_{tot} versus g_{fib} plot include the $g-r$ versus $r-z$ selection of the Main ELG cuts. They confirm the expectations that a total magnitude-based selection would add poor efficiency targets, likely extended targets with little flux inside

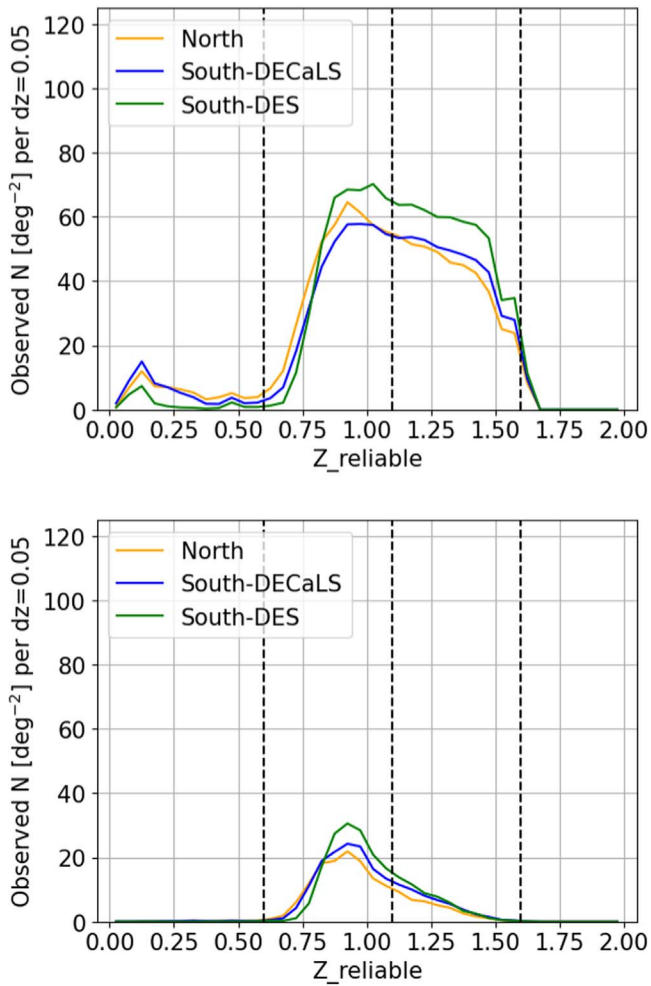


Figure 21. Redshift distributions for the Main ELG_LOP (top) and ELG_VLO (bottom) samples, split by footprint. The reported densities are the expected densities of the observed ELG targets providing a reliable z_{spec} (see the text for more details). The vertical dashed lines emphasize the relevant redshift ranges for the ELGs: $z = 0.6$, $z = 1.1$, and $z = 1.6$.

the DESI fibers. This justifies our choice of using a fiber magnitude-based cut for the Main selection.

7.7. Redshift Distribution and Total Numbers

We present the Main Survey ELG sample redshift distribution, along with the expected final densities and numbers of observed reliable redshifts.

To estimate these, we consider the Main Survey observations up to 2021 December, as this allows us to build a sample observed over a footprint that is fairly representative of the full DESI footprint (see Section 6.2). We restrict the sample to tiles with $800 \text{ s} < \text{EFFTIME_SPEC} < 1200 \text{ s}$. This sample contains about 1.5 million ELG_LOP spectra (North: 0.2 million; South-DECaLS: 1 million; and South-DES: 0.3 million) and about 187,000 ELG_VLO spectra (North: 21,000; South-DECaLS: 132,000; and South-DES: 34,000).

For each sample and each footprint, Table 4 lists the overall efficiency, i.e., the fraction of observed ELG spectra that provide a reliable z_{spec} (Equation (1)), and the efficiencies in the $0 < z < 0.6$, $0.6 < z < 1.1$, and $1.1 < z < 1.6$ ranges.

Figure 21 shows the expected density of the observed ELG targets providing a reliable z_{spec} at the end of the survey, when all the passes will have been completed. The distributions are

normalized to the target density (Table 1), multiplied by the fiber assignment rate at the end of the survey (Section 6.3), then multiplied by the fraction of observed ELG targets providing a reliable z_{spec} (Table 4). Altogether, the normalizations for the three footprints range from 910 to 1020 deg^{-2} for the ELG_LOP selection and from 160 to 210 deg^{-2} for the ELG_VLO selection. Table 5 reports the expected densities and total numbers. We caution that this is an estimate; for instance, the actual fiber assignment rate at the end of the survey could be different for operational reasons, or the redshift efficiency could be increased, with further work on the pipeline and analysis side.

7.8. ELG_LOP Sample

The ELG_LOP redshift distribution is expected to provide 400 deg^{-2} observed reliable z_{spec} in the $0.6 < z < 1.1$ range, and 460 deg^{-2} observed reliable z_{spec} in the $1.1 < z < 1.6$ range. It thus passes the DESI requirements, which are for 400 deg^{-2} in each redshift range (DESI Collaboration et al. 2023a, in preparation).

As expected, the deeper the imaging, the more reliable the z_{spec} that are gathered in the $1.1 < z < 1.6$ range. The North and South-DECaLS footprints show comparable $z < 0.6$ contamination, of 5%, but the South-DECaLS footprint provides 30 deg^{-2} more $z > 1.1$ redshifts. The South-DES footprint, with imaging 0.5 mag or more deeper than in South-DECaLS, has significantly better performances, with almost no $z < 0.6$ contamination. As it also benefits from a higher fiber assignment rate, the South-DES footprint should bring 40 deg^{-2} more $0.6 < z < 1.1$ and 80 deg^{-2} more $1.1 < z < 1.6$ reliable redshifts than the South-DECaLS footprint, despite its overall target density being smaller by 50 deg^{-2} .

Overall, we expect that 18.7 million ELG_LOP targets will be observed, with 12.9 million providing a reliable redshift, and 12.0 million providing a reliable redshift in $0.6 < z < 1.6$.

7.9. ELG_VLO Sample

The bottom panel of Figure 21 displays the redshift distribution of the ELG_VLO sample. We remind the reader that the fiber assignment rate for this sample is low (0.42, on average). As seen in Section 5.3, this selection is less sensitive to the imaging depth, leading to less significant differences in the redshift distribution among the three footprints. Nevertheless, their performances are still in the same order, the South-DES footprint providing the cleaner, higher-redshift sample, with the North footprint performing less well at high redshift.

For all footprints, the ELG_VLO selection should provide about 130 deg^{-2} reliable redshifts in the $0.6 < z < 1.1$ range and about 50 deg^{-2} reliable redshifts in the $1.1 < z < 1.6$ range.

Overall, we expect that 2.7 million ELG_VLO targets will be observed, with 2.5 million providing a reliable redshift and 2.4 million providing a reliable redshift in $0.6 < z < 1.6$.

7.10. Future Improvements

Characterizing the redshift distribution of the $\sim 30\%$ of ELG_LOP targets that do not provide a reliable z_{spec} will be an important topic for the DESI ELG LSS analysis, especially estimating the fraction and distribution in the redshift range of interest. This can be achieved by using, for instance, accurate z_{phot} or the clustering redshift method (e.g., Newman 2008).

Table 5
Expected Densities and Total Numbers for the ELG Samples at the End of the Main Survey

Sample	Footprint	Target		Observed		Observed with a Valid Redshift							
						All Redshifts		0 < z < 0.6		0.6 < z < 1.1		1.1 < z < 1.6	
		(deg ⁻²)		(deg ⁻²)		(deg ⁻²)		(deg ⁻²)		(deg ⁻²)		(deg ⁻²)	
ELG_LOP	North	1930	8.5M	1330	5.9M	950	4.2M	70	0.3M	440	1.9M	430	1.9M
	South-DECaLS	1950	16.6M	1350	11.4M	910	7.8M	70	0.6M	390	3.3M	460	3.9M
	South-DES	1900	2.1M	1390	1.5M	1010	1.1M	30	0.0M	430	0.5M	540	0.6M
	DESI	1940	27.2M	1340	18.7M	920	12.9M	70	0.9M	400	5.6M	460	6.4M
ELG_VLO	North	410	1.8M	180	0.8M	160	0.7M	0	0.0M	120	0.5M	40	0.2M
	South-DECaLS	490	4.2M	200	1.7M	190	1.6M	0	0.0M	130	1.1M	50	0.4M
	South-DES	480	0.5M	220	0.2M	210	0.2M	0	0.0M	150	0.2M	60	0.1M
	DESI	460	6.4M	190	2.7M	180	2.5M	0	0.0M	130	1.8M	50	0.6M

Note. Densities are converted to total numbers using the DESI areas reported in Table 1. The “Observed” column is the target column multiplied by the expected fiber assignment rate at the end of the survey (see Section 6.3). The “Observed with a Valid Redshift” column is the observed column multiplied by the efficiency reported in Table 4.

Besides, this fraction could be decreased thanks to further developments. First, it is very likely that some failures are spurious targets that are close to bright or medium stars, as the angular masking has purposely been chosen to be very minimal at the targeting step. Preparatory work has shown that the target selection indeed has some overdensity close to bright or medium stars. Another potential improvement could come from the pipeline, with, for instance, better sky subtraction. Last, as already mentioned, the reliable z_{spec} criterion could be refined. We will also run the QSO pipeline classification (Chaussidon et al. 2022) to identify secure QSO spectra; we expect that these could represent a non-negligible fraction of the failures of the Equation (1) criterion.

Nevertheless, we expect the fraction of redshift failures to remain non-negligible in the end. The VI analysis done on deep exposures has shown that about 14% of the ELG_LOP selection do not provide a VI-reliable z_{spec} ; even if the VI were to be done on data processed with a less advanced reduction pipeline, it would give the order of magnitude of the effect.

As presented in this paper, a specificity of the ELG sample is its variation across the footprints (North, South-DECaLS, and South-DES), in terms of the target density, redshift distribution, and redshift failures. For the cosmological analysis, while it will be possible to analyze these three footprints separately and then combine them (as was done for the eBOSS/ELG sample, for instance; see Raichoor et al. 2021), work is ongoing in the DESI collaboration to understand and model the variations at the target level, which would allow the whole footprint to be analyzed at once. Promising approaches include, for instance, injecting fake sources into the images and running the photometric pipeline (e.g., Kong et al. 2020), or identifying a subset of imaging properties and foregrounds, like the depths and Galactic extinction, which would allow the observed variations to be reproduced.

8. Conclusion

The ELG sample will constitute one-third of the 40 million extragalactic DESI redshifts and it will be used to probe the universe over the $0.6 < z < 1.6$ range, and in particular over the $1.1 < z < 1.6$ range, where it will bring the tightest of the DESI cosmological constraints. We have presented the DESI ELG target selection used for SV and the final selection that was

derived from it for the Main Survey. The Main Survey ELG selection comprises two disjoint sets of cuts, the ELG_LOP and the ELG_VLO selections, which have target densities of about 1940 deg^{-2} and 460 deg^{-2} , respectively. The ELG_LOP sample, which has a higher fiber assignment priority, favors the $1.1 < z < 1.6$ range, whereas the ELG_VLO sample, at lower fiber assignment priority, favors the $0.6 < z < 1.1$ range. These two samples are completed by the ELG_HIP sample, a random 10% subsample of the ELG_LOP and ELG_VLO samples, which has the same fiber assignment priority as the LRG one.

The target selection is based on the grz -band photometry from LS-DR9. As the ELG targets are at low S/N in the imaging, we define three footprints, isolating the three regions linked to the underlying observing programs with different imaging depths: North, South-DECaLS, and South-DES. Both the Main Survey ELG_LOP and ELG_VLO samples are selected with a g -band fiber magnitude cut $g_{\text{fib}} < 24.1$, which favors [O II] emitters and more successful z_{spec} measurement rates, as well as a specific $(g-r)$ versus $(r-z)$ color box, which primarily selects the redshift range. The SV ELG sample, which was used to tune the Main Survey cuts, is an extended version of these, noticeably toward: (1) bluer $r-z$ targets, exploring the $1.1 < z < 1.6$ range; (2) redder $r-z$ targets, exploring the $0.6 < z < 1.1$ range; (3) fainter blue targets; and (4) g -band total magnitude-selected targets.

We then present the photometric properties of the Main Survey ELG selection. In terms of the magnitude, the ELG_LOP sample is 0.2 mag fainter in the r band and 0.5 mag fainter in the z band than the ELG_VLO sample, due to the different selection boxes in the $g-r$ versus $r-z$ diagram. For both samples, the target density is slightly different in each footprint, mostly because of the difference in imaging depth: ELG_LOP ranges from 1900 to 1950 deg^{-2} and ELG_VLO ranges from 410 to 490 deg^{-2} . The imaging and foreground maps causing the largest target density fluctuations are the imaging depth, in particular in the g band, and the Galactic dust extinction. Attempts to correct these fluctuations have been deferred to subsequent papers, when a cleaner sample will have been defined, e.g., after the inclusion of appropriate angular masking and correction for z_{spec} measurement failures.

Last, we present the spectroscopic properties of the Main Survey ELG selection. For this purpose, we use observations from three DESI surveys, covering two phases of validation and the first seven months of the Main Survey. We define a

preliminary criterion for selecting reliable z_{spec} measurements, which requires a minimal [O II] doublet flux S/N as a function of the χ^2 difference between the first and second redshift values that best fit the observed spectrum. This criterion exploits the fact that the [O II] doublet is the key emission line for measuring accurate redshifts of star-forming ELGs. Using VI spectra tiles and repeat observations, we demonstrate that such a criterion is extremely efficient, since it selects most of the VI-confirmed z_{spec} and keeps the fraction of catastrophic z_{spec} measurements below 1%. Nevertheless, this discards about 30% of the observed ELG_LOP spectra (and 6% of the ELG_VLO ones): even if some improvements in the data reduction and the reliability criterion could reduce those percentages, we expect that it will remain non-negligible for the ELG_LOP sample, and it will be necessary to characterize the redshift properties of the discarded spectra.

We define the efficiency in a given redshift range as the fraction of observed ELG spectra providing a reliable z_{spec} in that redshift range. Depending on the footprint, the ELG_LOP selection has efficiencies of 2%–5%, 29%–33%, and 32%–39% in the $0 < z < 0.6$, $0.6 < z < 1.1$, and $1.1 < z < 1.6$ ranges, respectively, with deeper imaging providing less contamination and more high-redshift spectra. This sample will thus fulfill the DESI requirements: with the expected fiber assignment rate of 0.69, it should provide 400 deg^{-2} and 460 deg^{-2} observed reliable z_{spec} in the $0.6 < z < 1.1$ and $1.1 < z < 1.6$ ranges, respectively. Overall, we expect that 18.7 million ELG_LOP targets will be observed, with 12.9 million providing a reliable redshift, and 12.0 million providing a reliable redshift in $0.6 < z < 1.6$.

The ELG_VLO selection has efficiencies of 0%–1%, 67%–70%, and 21%–26% in the $0 < z < 0.6$, $0.6 < z < 1.1$, and $1.1 < z < 1.6$ ranges, respectively. As expected from its design, the sample has a very high overall efficiency, extremely few contaminants, and it peaks in the $0.6 < z < 1.1$ range. With the expected fiber assignment rate of 0.42, that should provide 130 deg^{-2} and 50 deg^{-2} reliable z_{spec} in the $0.6 < z < 1.1$ and $1.1 < z < 1.6$ ranges, respectively. Overall, we expect that 2.7 million ELG_VLO targets will be observed, with 2.5 million providing a reliable redshift and 2.4 million providing a reliable redshift in $0.6 < z < 1.6$.

J.M. gratefully acknowledges support from the U.S. Department of Energy, Office of Science, Office of High Energy Physics, under Award Number DE-SC0020086.

This research is supported by the Director, Office of Science, Office of High Energy Physics of the U.S. Department of Energy, under Contract No. DE-AC02-05CH11231, and by the National Energy Research Scientific Computing Center, a DOE Office of Science User Facility, under the same contract; additional support for DESI is provided by the U.S. National Science Foundation, Division of Astronomical Sciences, under Contract No. AST-0950945 to the NSF’s National Optical-Infrared Astronomy Research Laboratory; the Science and Technologies Facilities Council of the United Kingdom; the Gordon and Betty Moore Foundation; the Heising-Simons Foundation; the French Alternative Energies and Atomic Energy Commission (CEA); the National Council of Science and Technology of Mexico (CONACYT); the Ministry of Science and Innovation of Spain (MICINN), and the DESI Member Institutions: <https://www.desi.lbl.gov/collaborating-institutions>.

The DESI Legacy Imaging Surveys consist of three individual and complementary projects: the Dark Energy Camera Legacy Survey (DECaLS), the Beijing–Arizona Sky Survey (BASS), and the Mayall z-band Legacy Survey (MzLS). DECaLS, BASS, and MzLS together include data obtained, respectively, at the Blanco telescope, Cerro Tololo Inter-American Observatory, NSF’s NOIRLab; the Bok telescope, Steward Observatory, University of Arizona; and the Mayall telescope, Kitt Peak National Observatory, NOIRLab. NOIRLab is operated by the Association of Universities for Research in Astronomy (AURA) under a cooperative agreement with the National Science Foundation. Pipeline processing and analyses of the data were supported by NOIRLab and the Lawrence Berkeley National Laboratory. Legacy Surveys also use data products from the Near-Earth Object Wide-field Infrared Survey Explorer (NEOWISE), a project of the Jet Propulsion Laboratory/California Institute of Technology, funded by the National Aeronautics and Space Administration. Legacy Surveys are supported by: the Director, Office of Science, Office of High Energy Physics of the U.S. Department of Energy; the National Energy Research Scientific Computing Center, a DOE Office of Science User Facility; the U.S. National Science Foundation, Division of Astronomical Sciences; the National Astronomical Observatories of China, the Chinese Academy of Sciences; and the Chinese National Natural Science Foundation. LBNL is managed by the Regents of the University of California, under contract to the U.S. Department of Energy. The complete acknowledgments can be found at <https://www.legacysurvey.org>.

The authors are honored to be permitted to conduct scientific research on Iolkam Du’ag (Kitt Peak), a mountain with particular significance to the Tohono O’odham Nation.

We thank the anonymous referee for the constructive comments, which helped us to improve the clarity of the paper.

Facility: Mayall.

Data Availability

The ELG targets for the SV, the One-Percent, and the Main Surveys are accessible at <https://data.desi.lbl.gov/public/ets/target/>. We refer the reader to Myers et al. (2022) for a description of the files and the structure of the folders.

The spectroscopic data for the ELG targets from the SV and One-Percent programs—along with other DESI tracers—will be part of the DESI Early Data Release, which is tentatively scheduled for the first half of 2023. The Main Survey spectroscopic data will be part of the DESI DR1 release, which will be later.

All the data points shown in the published graphs are available in a machine-readable form on the following website: <https://doi.org/10.5281/zenodo.6950999>.

Appendix A

SV Survey Target Selection Cuts

We present in Table A1 the detailed cuts of the SV Survey ELG selection, discussed in Section 4. This selection comprises the union of two samples, SVGTOT and SVGFIB, which have similar cuts, but are based either on g_{tot} or g_{fib} . The overall selection has a target density of $\sim 7000 \text{ deg}^{-2}$, as the SVGFIB and SVGTOT selections have a large overlap.

The names SVGTOT and SVGFIB are names that are assigned to targeting bits by `desitarget`, the target selection

Table A1
SV Survey Target Selection Cuts

Sample	Density	Cuts	Comment
Clean	...	brick_primary=True nobs_{grz} > 0 $\text{flux}_{\{grz\}} \times \sqrt{\text{flux_ivar}_{\{grz\}}} > 0$ (maskbits & 2 ¹) = 0, (maskbits & 2 ¹²) = 0, (maskbits & 2 ¹³) = 0	Unique object Observed in the grz bands Positive S/N in the grz bands Not close to bright star/galaxy
SVGTOT	~5200 deg ⁻²	Clean g > 20 $(g - r) + 1.2 \times (r - z) < 1.6 - 7.2 \times (g_{\text{tot}} - \text{GTOTFAINT_FDR})$ (g - r < 0.2) or (g - r < 1.15 × (r - z) + LOWZCUT_ZP + 0.10) g - r < -1.2 × (r - z) + 2.0	Clean sample Bright cut Sliding faint cut Star/low-z cut Redshift/[O II] cut
SVGFB	~5600 deg ⁻²	Clean g > 20 $(g - r) + 1.2 \times (r - z) < 1.6 - 7.2 \times (g_{\text{fib}} - \text{GFIBFAINT_FDR})$ (g - r < 0.2) or (g - r < 1.15 × (r - z) + LOWZCUT_ZP + 0.10) g - r < -1.2 × (r - z) + 2.0	Clean sample Bright cut Sliding faint cut Star/low-z cut Redshift/[O II] cut
FDRGTOT	~2400 deg ⁻²	Clean g > 20 g < GTOTFAINT_FDR 0.3 < r - z < 1.6 g - r < 1.15 × (r - z) + LOWZCUT_ZP g - r < -1.2 × (r - z) + 1.6	Clean sample Bright cut Faint cut r - z cut Star/low-z cut Redshift/[O II] cut
FDRGFIB	~2500 deg ⁻²	Clean g > 20 g _{fib} < GFIBFAINT_FDR 0.3 < r - z < 1.6 g - r < 1.15 × (r - z) + LOWZCUT_ZP g - r < -1.2 × (r - z) + 1.6	Clean sample Bright cut Faint cut r - z cut Star/low-z cut Redshift/[O II] cut

Note. The following quantities have values defined for the North and South regions: GTOTFAINT_FDR: North = 23.5, South = 23.4; GFIBFAINT_FDR: North = 24.1, South = 24.1; and LOWZCUT_ZP: North = -0.20, South = -0.15. See Table 2 for the definitions of the terms in the cuts.

pipeline (Myers et al. 2022). For completeness, we also report the cuts for the FDRGTOT and FDRGFIB selections, which are other targeting bits. The FDRGTOT (FDRGFIB) selection is fully included in the SVGTOT (SVGFB) sample.

Appendix B HSC/DR2 z_{phot} Comparison to DESI/ELG z_{spec}

We illustrate in Figure B1 how the HSC/DR2 z_{phot} estimates perform against the Main Survey ELG DESI z_{spec} measurements corresponding to the observations until 2021 December (see Section 6.2). We use the DEMO code z_{phot} ; we emphasize that DESI z_{spec} data were not used to train the HSC/DR2 z_{phot} method. The matched sample has ~175,000 targets with a z_{spec} passing our reliability criterion of Equation (1). We do not make any quality cuts on the HSC z_{phot} .

The left panels compare the two redshift measurements, with the color-coding indicating the HSC risk parameter, which quantifies the reliability of the z_{phot} estimate (with $\text{risk}=0$ being the most secure and $\text{risk}=1$ the least secure; see Tanaka et al. 2018). Overall, the HSC z_{phot} estimates perform very well, noticeably over the whole $0.6 < z < 1.6$ range, thanks to its very deep imaging and the presence of y-band imaging. This justifies the use of the HSC z_{phot} in Figures 3 and

4. In particular, the risk parameter provides a sensible estimation of the z_{phot} reliability. This last point is illustrated in the bottom panel on the left, where we also display the median value of $\Delta z = (z_{\text{phot}} - z_{\text{spec}})/(1 + z_{\text{spec}})$ for three subsamples: all matches (black), the 50% lower risk parameter values (cyan), and the 25% lower risk parameter values (red). The $1.48 \times \text{MAD}(\Delta z)$ values are displayed as shaded regions, with typical values of 0.08 (all matches), 0.05 (50% lower risk), and 0.02 (25% lower risk).

Nevertheless, the cuts on risk are biasing the sample toward the redder ELG targets—hence the lower-redshift ones, as those are easier to model with the z_{phot} algorithms; said differently, applying cuts on risk will exclude the blue high-redshift ELGs from the sample. This is illustrated in the right panels, where we plot for each subsample (all, 50%, and 25% lower risk parameter values) the DESI z_{spec} distribution (filled histograms) and the HSC z_{phot} distribution (empty histograms); the DESI z_{spec} distribution with no cut on risk is displayed via the black hatched histogram. As a consequence, a careful balance between the z_{phot} reliability and the sample representativeness will be required for any analysis using the HSC z_{phot} distribution for an ELG DESI-like sample, as, for instance, when trying to infer the redshift distribution of the DESI ELG targets that do not provide a reliable z_{spec} .

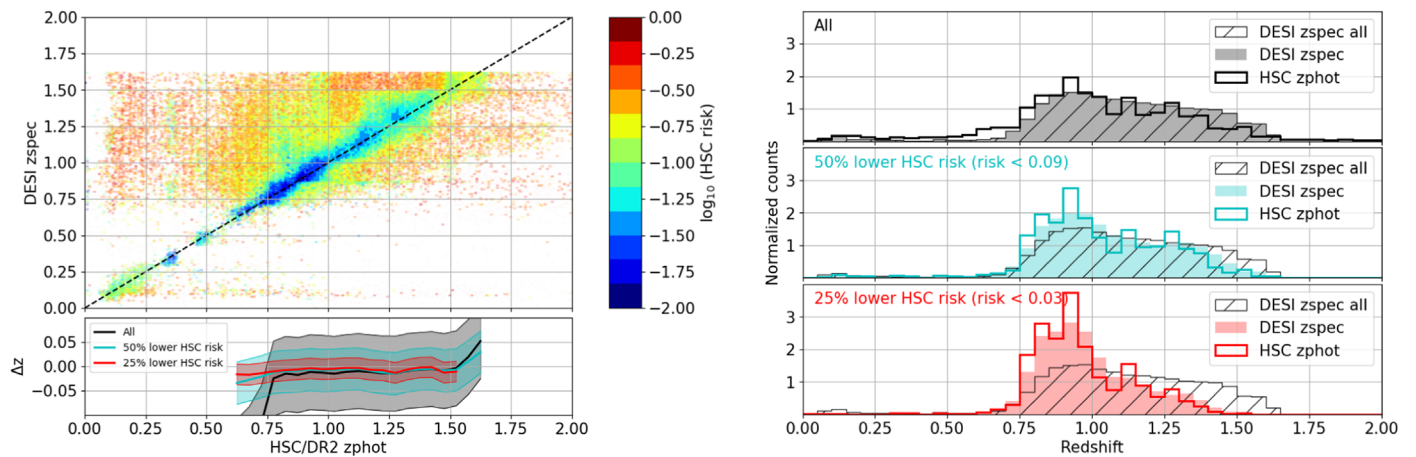


Figure B1. Comparison of the HSC/DR2 z_{phot} to the DESI Main Survey ELG z_{spec} measurements, using $\sim 175,000$ matched targets with a z_{spec} passing our reliability criterion of Equation (1). Left: scatter plots color-coded according to the average HSC risk parameter (see the text); the solid lines in the bottom panel are the median values of $\Delta z = (z_{\text{phot}} - z_{\text{spec}})/(1 + z_{\text{spec}})$ for three subsamples—all the matches (black), the 50% lower risk parameter values (cyan), and the 25% lower risk parameter values (red), with the shaded regions showing the $1.48 \times \text{MAD}(\Delta z)$ values. Right: redshift distributions for these three subsamples; in each panel, the filled (empty) histogram is for the DESI z_{spec} (HSC z_{phot}). The DESI z_{spec} distribution with no cut on risk is displayed via the black hatched histogram.

ORCID iDs

A. Raichoor <https://orcid.org/0000-0001-5999-7923>
 J. Moustakas <https://orcid.org/0000-0002-2733-4559>
 Jeffrey A. Newman <https://orcid.org/0000-0001-8684-2222>
 T. Karim <https://orcid.org/0000-0002-5652-8870>
 S. Ahlen <https://orcid.org/0000-0001-6098-7247>
 Shadab Alam <https://orcid.org/0000-0002-3757-6359>
 S. Bailey <https://orcid.org/0000-0003-4162-6619>
 D. Brooks <https://orcid.org/0000-0002-8458-5047>
 K. Dawson <https://orcid.org/0000-0002-0553-3805>
 A. Dey <https://orcid.org/0000-0002-4928-4003>
 Biprateep Dey <https://orcid.org/0000-0002-5665-7912>
 G. Dhungana <https://orcid.org/0000-0002-5402-1216>
 D. J. Eisenstein <https://orcid.org/0000-0002-2929-3121>
 A. Font-Ribera <https://orcid.org/0000-0002-3033-7312>
 J. García-Bellido <https://orcid.org/0000-0002-9370-8360>
 E. Gaztañaga <https://orcid.org/0000-0001-9632-0815>
 S. Gontcho A Gontcho <https://orcid.org/0000-0003-3142-233X>
 J. Guy <https://orcid.org/0000-0001-9822-6793>
 M. Ishak <https://orcid.org/0000-0002-6024-466X>
 R. Kehoe <https://orcid.org/0000-0002-7101-697X>
 T. Kisner <https://orcid.org/0000-0003-3510-7134>
 Anthony Kremin <https://orcid.org/0000-0001-6356-7424>
 Ting-Wen Lan <https://orcid.org/0000-0001-8857-7020>
 M. Landriau <https://orcid.org/0000-0003-1838-8528>
 L. Le Guillou <https://orcid.org/0000-0001-7178-8868>
 Michael E. Levi <https://orcid.org/0000-0003-1887-1018>
 P. Martini <https://orcid.org/0000-0002-4279-4182>
 Aaron M. Meisner <https://orcid.org/0000-0002-1125-7384>
 Jundan Nie <https://orcid.org/0000-0001-6590-8122>
 N. Palanque-Delabrouille <https://orcid.org/0000-0003-3188-784X>
 W. J. Percival <https://orcid.org/0000-0002-0644-5727>
 F. Prada <https://orcid.org/0000-0001-7145-8674>
 C. G. Sabiu <https://orcid.org/0000-0002-5513-5303>
 E. F. Schlafly <https://orcid.org/0000-0002-3569-7421>
 D. Schlegel <https://orcid.org/0000-0002-5042-5088>
 Gregory Tarlé <https://orcid.org/0000-0003-1704-0781>
 Christophe Yèche <https://orcid.org/0000-0001-5146-8533>

Rongpu Zhou <https://orcid.org/0000-0001-5381-4372>

Zhimin Zhou <https://orcid.org/0000-0002-4135-0977>

H. Zou <https://orcid.org/0000-0002-6684-3997>

References

- Aihara, H., Arimoto, N., Armstrong, R., et al. 2018, *PASJ*, 70, S4
 Aihara, H., AlSayyad, Y., Ando, M., et al. 2019, *PASJ*, 71, 114
 Alam, S., Aubert, M., Avila, S., et al. 2021, *PhRvD*, 103, 083533
 Alexander, D. M., Davis, T. M., Chaussidon, E., et al. 2022, arXiv:2208.08517
 Allende Prieto, C., Cooper, A. P., Dey, A., et al. 2020, *MNRAS*, 4, 188
 Bianchi, D., Burden, A., Percival, W. J., et al. 2018, *MNRAS*, 481, 2338
 Bruzual, G., & Charlot, S. 2003, *MNRAS*, 344, 1000
 Chambers, K. C., Magnier, E. A., Metcalfe, N., et al. 2016, arXiv:1612.05560
 Chaussidon, E., Yèche, C., Palanque-Delabrouille, N., et al. 2022, arXiv:2208.08511
 Colless, M., Peterson, B. A., Jackson, C., et al. 2003, arXiv:0306581
 Comparat, J., Delubac, T., Jouvel, S., et al. 2016, *A&A*, 592, A121
 Cooper, A. P., Kogosov, S. E., Allende Prieto, C., et al. 2022, arXiv:2208.08514
 Dawson, K. S., Schlegel, D. J., Ahn, C. P., et al. 2013, *AJ*, 145, 10
 Dawson, K. S., Kneib, J.-P., Percival, W. J., et al. 2016, *AJ*, 151, 44
 DESI Collaboration, Abareshi, B., Aguilar, J., et al. 2022, *AJ*, 164, 207
 DESI Collaboration, Aghamousa, A., Aguilar, J., et al. 2016a, arXiv:1611.00036
 DESI Collaboration, Aghamousa, A., Aguilar, J., et al. 2016b, arXiv:1611.00037
 Dey, A., Schlegel, D. J., Lang, D., et al. 2019, *AJ*, 157, 168
 Drinkwater, M. J., Jurek, R. J., Blake, C., et al. 2010, *MNRAS*, 401, 1429
 Eisenstein, D. J., & Hu, W. 1998, *ApJ*, 496, 605
 Eisenstein, D. J., Zehavi, I., Hogg, D. W., et al. 2005, *ApJ*, 633, 560
 Flaugher, B., Diehl, H. T., Honscheid, K., et al. 2015, *AJ*, 150, 150
 Gaia Collaboration, Brown, A. G. A., Vallenari, A., et al. 2018, *A&A*, 616, A1
 Górski, K. M., Hivon, E., Banday, A. J., et al. 2005, *ApJ*, 622, 759
 Guy, J., Bailey, S., Kremin, A., et al. 2022, arXiv:2209.1448
 Hahn, C., Wilson, M. J., Ruiz-Macias, O., et al. 2022, arXiv:2208.08512
 Ivezić, Ž., Kahn, S. M., Tyson, J. A., et al. 2019, *ApJ*, 873, 111
 Jones, D. H., Read, M. A., Saunders, W., et al. 2009, *MNRAS*, 399, 683
 Kaiser, N. 1987, *MNRAS*, 227, 1
 Karim, T., Lee, J. H., Eisenstein, D. J., et al. 2020, *MNRAS*, 497, 4587
 Kong, H., Burleigh, K. J., Ross, A., et al. 2020, *MNRAS*, 499, 3943
 Lan, T.-W., Tojeiro, R., Armengaud, E., et al. 2023, *ApJ*, 943, 68
 Lang, D., Hogg, D. W., & Mykytyn, D. 2016, The Tractor: Probabilistic Astronomical Source Detection and Measurement, Astrophysics Source Code Library, ascl:1604.008
 Laureijs, R., Amiaux, J., Arduini, S., et al. 2011, arXiv:1110.3193
 Le Fèvre, O., Cassata, P., Cucciati, O., et al. 2013, *A&A*, 559, A14
 Levi, M., Bebek, C., Beers, T., et al. 2013, arXiv:1308.0847
 Madau, P., & Dickinson, M. 2014, *ARA&A*, 52, 415

- Mancone, C., & Gonzalez, A. 2012, EzGal: A Flexible Interface for Stellar Population Synthesis Models, Astrophysics Source Code Library, ascl:1208.021
- Meisner, A. M., Lang, D., Schlafly, E. F., & Schlegel, D. J. 2021, *RNAAS*, **5**, 200
- Mohammad, F. G., Percival, W. J., Seo, H.-J., et al. 2020, *MNRAS*, **498**, 128
- Moustakas, J., Kennicutt, R. C., Jr., Tremonti, C. A., et al. 2006, *ApJ*, **642**, 775
- Myers, A. D., Palanque-Delabrouille, N., Prakash, A., et al. 2015, *ApJS*, **221**, 27
- Myers, A. D., Moustakas, J., Bailey, S., et al. 2023, *AJ*, **165**, 50
- Newman, J. A. 2008, *ApJ*, **684**, 88
- Newman, J. A., Cooper, M. C., Davis, M., et al. 2013, *ApJS*, **208**, 5
- Oke, J. B., & Gunn, J. E. 1983, *ApJ*, **266**, 713
- Padmanabhan, N., Schlegel, D. J., Finkbeiner, D. P., et al. 2008, *ApJ*, **674**, 1217
- Perlmutter, S., Aldering, G., Goldhaber, G., et al. 1999, *ApJ*, **517**, 565
- Raichoor, A., Comparat, J., Delubac, T., et al. 2017, *MNRAS*, **471**, 3955
- Raichoor, A., Eisenstein, D. J., Karim, T., et al. 2020, *RNAAS*, **4**, 180
- Raichoor, A., deMattia, A., Ross, A. J., et al. 2021, *MNRAS*, **500**, 3254
- Riess, A. G., Filippenko, A. V., Challis, P., et al. 1998, *AJ*, **116**, 1009
- Ruiz-Macias, O., Zarrouk, P., Cole, S., et al. 2020, *RNAAS*, **4**, 187
- Schlegel, D. J., Finkbeiner, D. P., & Davis, M. 1998, *ApJ*, **500**, 525
- Silber, J. H., Fagrelus, P., Fanning, K., et al. 2023, *AJ*, **165**, 9
- Takada, M., Ellis, R. S., Chiba, M., et al. 2014, *PASJ*, **66**, R1
- Tanaka, M., Coupon, J., Hsieh, B.-C., et al. 2018, *PASJ*, **70**, S9
- The Dark Energy Survey Collaboration 2005, arXiv:0510346
- Weinberg, D. H., Mortonson, M. J., Eisenstein, D. J., et al. 2013, *PhR*, **530**, 87
- Yèche, C., Palanque-Delabrouille, N., Claveau, C.-A., et al. 2020, *RNAAS*, **4**, 179
- York, D. G., Adelman, J., Anderson, J. E., Jr., et al. 2000, *AJ*, **120**, 1579
- Zhou, R., Newman, J. A., Dawson, K. S., et al. 2020, *RNAAS*, **4**, 181
- Zhou, R., Dey, B., Newman, J. A., et al. 2023, *AJ*, **165**, 58
- Zou, H., Zhou, X., Fan, X., et al. 2017, *PASP*, **129**, 064101