Improved Measurements of Galaxy Star Formation Stochasticity from the Intrinsic Scatter of Burst Indicators

Adam Broussard,1 Eric Gawiser,1 and Kartheik Iyer2

1Department of Physics and Astronomy, Rutgers, The State University of New Jersey, 136 Frelinghuysen Rd, Piscataway, NJ 08854, USA
2The Dunlap Institute for Astronomy and Astrophysics, 50 St. George Street, Toronto, Ontario, Canada M5S 3H4

(Dated: June 2022, Submitted to ApJ)

ABSTRACT

Measurements of short-timescale star formation variations (i.e., “burstiness”) are integral to our understanding of star formation feedback mechanisms and the assembly of stellar populations in galaxies. We expand upon the work of Broussard et al. (2019) by introducing a new analysis of galaxy star formation burstiness that accounts for variations in the $Q_{sg} = E(B-V)_{stars}/E(B-V)_{gas}$ distribution, a major confounding factor. We use Balmer decrements from the MOSFIRE Deep Evolution Field (MOSDEF) survey to measure $Q_{sg}$, which we use to construct mock catalogs from the Santa Cruz Semi-Analytic Models and MUFASA cosmological hydrodynamical simulation based on 3D-HST, Fiber Multi-Object Spectrograph (FMOS)-COSMOS, and MOSDEF galaxies with Hα detections. The results of the mock catalogs are compared against observations using the burst indicator $\eta = \log_{10}(SFR_{H\alpha}/SFR_{NUV})$, with the standard deviation of the $\eta$ distribution indicating burstiness. We find decent agreement between mock and observed $\eta$ distribution shapes; however, the FMOS-COSMOS and MOSDEF mocks show a systematically low median and scatter in $\eta$ in comparison to the observations. This work also presents the novel approach of analytically deriving the relationship between the intrinsic scatter in $\eta$, scatter added by measurement uncertainties, and observed scatter, resulting in an intrinsic burstiness measurement of $0.06 - 0.16$ dex.

Keywords: Galaxy Evolution, Star Formation, Starburst Galaxies

1. INTRODUCTION

Star formation is a key component of the formation and evolution of galaxies. Galaxy star formation histories (SFHs) bear imprints of the various physical processes that regulate their star formation across a broad range of timescales. Some examples of these processes and their relevant timescales include supermassive black hole feedback (Anglés-Alcázar et al. 2017, ; $\gtrsim 1$ Gyr), inflows and outflows of gas (Pillepich et al. 2017; Weinberger et al. 2017, ; $100 - 500$ Myr), galaxy mergers (Davé et al. 2017, ; $10 - 100$ Myr), and supernova feedback (Körtgen et al. 2016, ; $1 - 20$ Myr). An understanding of the variation in galaxy SFHs is central to understanding the role these processes play in galaxy evolution.

Analyses of star formation “burstiness” focus on short-timescale variations ($\lesssim 100$ Myr), which dominate the evolution of dwarf galaxies in simulations (Shen et al. 2014; Domínguez et al. 2015; Tacchella et al. 2022) and in observations of nearby galaxies (Kauffmann 2014; Weisz et al. 2014; Benítez-Llambay et al. 2015). Burstiness also affects observed correlations such as the extensively-studied SFR-M$_*$ correlation (Daddi et al. 2007; Noeske et al. 2007; Salim et al. 2007; Wuyts et al. 2011; Kurczynski et al. 2016) by adding scatter. In particular, increased scatter has been predicted in simulations for dwarf galaxies (Somerville et al. 2015). Additionally, Faucher-Giguère (2018) has predicted a tran-
sition from ubiquitous bursty star formation in galaxies at high redshift to bursty star formation in low-mass galaxies and quiescent star formation in high-mass galaxies at low redshift.

Direct measurements of burstiness have been the subject of multiple recent papers for both observed (Shen et al. 2014; Dominguez et al. 2015; Guo et al. 2016; Sparre et al. 2017; Broussard et al. 2019; Wang & Lilly 2020; Emami et al. 2021) and simulated (Hopkins et al. 2014; Sparre et al. 2015) galaxies. Many of these use the Hα nebular emission line flux and the stellar ultraviolet continuum as tracers of recent star formation due to their sensitivity to stars formed in the past $\sim 5$ Myr and $\sim 200$ Myr respectively. Weisz et al. (2012) took the ratio of these two fluxes, finding that a toy model of periodically repeating star formation bursts produced flux ratio distributions consistent with those of 185 Spitzer-observed galaxies. Guo et al. (2016) found a decrease in the average Hβ-to-far-UV SFR ratio with decreasing M∗, noting burstiness as a plausible explanation for this variation and that, with decreasing M∗, burstiness is increasingly necessary to reproduce the observed level of variation. Wang & Lilly (2020) measured the power spectrum distribution (PSD) of star formation variability on various timescales, finding a power law slopes between 1.0 and 2.0 as well as a close relationship between specific star formation rate (sSFR) PSDs and the effective gas depletion timescale.

Broussard et al. (2019) showed that it is possible to estimate burstiness in a way that is robust to variations in the high-mass slope of the stellar IMF, metallicity, and (when assuming a Calzetti et al. 2000 dust law) dust measurement errors for 3D-HST galaxies at $z \sim 1$. Because it is difficult to measure recent variations in a single galaxy’s star formation history, burstiness is instead measured by comparing short- ($\lesssim 10$ Myr) and long-timescale ($\sim 100$ Myr) star formation tracers using a burst indicator $\eta = \log_{10} (\text{SFR}_{\text{Hα}}/\text{SFR}_{\text{NUV}})$, which yields positive values in the case of rising star formation rates, negative values for falling star formation rates, and $\sim 0$ for roughly constant star formation rates. Consequently, the distribution of burst indicators for an ensemble of galaxies is used to characterize the population burstiness, with the width of the distribution being the primary indicator of bursty star formation.

Despite efforts to account for various survey selection effects when generating mock catalogs from simulations, Broussard et al. (2019) noted a discrepancy in the relationship between $\eta$ and M∗ when comparing the mock catalogs and 3D-HST observations. The observed sample showed a strong positive trend for $\eta$ with increasing M∗ that was not present in the mocks. Tuning the ratio of stellar to nebular dust reddening $Q_{\text{sg}} = E(B-V)_{\text{stars}}/E(B-V)_{\text{gas}}$ was only able to partially resolve the discrepancy, as modifying $Q_{\text{sg}}$ from its typical low-redshift value of 0.44 (Calzetti et al. 2000) to $\sim 1$ removed the correlation of $\eta$ with M∗, but also necessitated a negative average value of $\eta$. This would have the unlikely implication of strongly declining SFHs for the selected sample of star forming galaxies. This motivates an independent treatment of the nebular and stellar dust in order to better understand burstiness at $z \gtrsim 1$.

While variations in $Q_{\text{sg}}$ add systematic scatter to the distribution of galaxy $\eta$ values and therefore bias measurements of burstiness, previous burstiness analyses have not accounted for this effect. In part, this is due to the rarity of observations that enable simultaneous independent dust measurements for each of the star formation tracers; however the relationship between E(B−V)stars and E(B−V)gas has been studied for a variety of dust curves outside of the context of burstiness (Calzetti et al. 2000; Kashino et al. 2013; Price et al. 2014; Reddy et al. 2015; Shivaei et al. 2016; Reddy et al. 2020; Shivaei et al. 2020). This work presents the first analysis to apply measurements of the $Q_{\text{sg}}$ distribution to burstiness analyses. Section 2 introduces the three surveys and two simulations we incorporate into our analysis. Section 3 describes the sample selection and methodology of measuring $Q_{\text{sg}}$. Our analysis consists of two approaches to measuring burstiness. The first, described in Section 4 involves the generation of mock catalogs from the MUFASA cosmological hydrodynamical simulation and Santa Cruz SAM that imitate the observational uncertainties and selection effects of our three observed surveys and therefore compare scatter in simulations against observations. The second, described in Section 5 introduces a novel burstiness estimation technique in which we subtract away variance in $\eta$ caused by observational uncertainties and $Q_{\text{sg}}$ variations to measure the underlying intrinsic $\eta$ scatter. These are followed by Section 6, in which we describe the results of our two-pronged analysis. We then discuss our findings in detail in Section 7 and present our final conclusions in Section 8.

2. DATA

In addition to the 3D-HST (Skelton et al. 2014) data set described in Broussard et al. (2019), we incorporate the FMOS-COSMOS and MOSDEF surveys as described below.

2.1. 3D-HST Observations

While we include a brief summary of the selections we implement for the 3D-HST catalog here to form the sample used throughout this work, a detailed description can be found in Broussard et al. (2019). This sample is selected to
Figure 1. Scatter plot of $E(B - V)_{\text{gas}}$ vs. $E(B - V)_{\text{stars}}$ for MOSDEF galaxies assuming a Calzetti et al. (2000) dust law. Dashed lines indicate various values of $Q_{sg} = E(B - V)_{\text{stars}}/E(B - V)_{\text{gas}}$. Many objects have $E(B - V)_{\text{gas}} < 0$, caused by emission line ratios with $F_{\text{H}\alpha}/F_{\text{H}\beta} < 2.86$. This indicates an unphysical inversion of the typical Balmer decrement dust correction.

have SNR$_{\text{H}\alpha} > 10$ to ensure acceptable data quality as determined by visual inspection of grism data for representative galaxies in the catalog. We also restrict the redshift range to $0.65 < z < 1.50$ for consistency with detections of H$\alpha$ using the G141 grism. Following Broussard et al. (2019), we estimate an AGN contamination rate of $\sim 3\%$ for our sample, which does not greatly affect our results, particularly because galaxies rejected for AGN contamination were not outliers in mass or SFR. A constant [NII] fraction of 0.1 was assumed for all galaxies when calculating H$\alpha$ fluxes, and the dust correction was calculated from SED fits using a Calzetti et al. (2000) dust curve, which uses $E(B - V)_{\text{stars}} = 0.44 E(B - V)_{\text{gas}}$. 3D-HST grism data were obtained by using a detection in the F 140W as a direct image, which serves to establish the calibration wavelength of the spectra. Detections in the G141 grism have an average 5$\sigma$ sensitivity of $f = 5.5 \times 10^{-17}$ erg s$^{-1}$ cm$^{-2}$ (Brammer et al. 2012).
2.2. FMOS-COSMOS Observations

A second source of spectroscopic data comes from the FMOS-COSMOS survey (Silverman et al. 2015). The Fiber Multi-Object Spectrograph (FMOS) is an instrument on the Subaru telescope in Mauna Kea, Hawaii. The purpose of the FMOS-COSMOS survey is to provide high-resolution spectra of \( \sim 1000 \) star forming galaxies at \( z \sim 1.5 \). The H-long grating covers the wavelength range \( 1.6 < \lambda < 1.8 \mu m \) enabling detections of H\( \alpha \) while followup observations with the J-long grating (\( 1.11 < \lambda < 1.35 \mu m \)) adds coverage of H\( \alpha \) and \([\text{OIII}]\lambda4959,5007\). Integration times enable detections of emission line fluxes with \( f = 4 \times 10^{-17} \text{ erg s}^{-1} \text{ cm}^{-2} \) at SNR = 3. In order to alleviate sample selection bias against dust-obscured galaxies, Herschel detections were used to select galaxies that lie on and off the star forming sequence. AGN are identified with X-ray confirmation using Chandra (Civano et al. 2012) and excluded from our final sample.

2.3. MOSDEF Observations

We also use spectroscopy from the MOSFIRE Deep Evolution Field (MOSDEF; Kriek et al. (2015)) survey. MOSDEF is a 47-night survey undertaken on the Keck MOSFIRE Spectrograph (McLean et al. 2010, 2012) that covers \( \sim 300 \) sq. arcmin. across the AEGIS (Davis et al. 2007), COSMOS (Scoville et al. 2007), and GOODS-N (Giavalisco et al. 2004) fields for the low-redshift regime of the survey (\( 1.37 < z < 1.70 \)), which is most comparable to 3D-HST. MOSDEF excludes AGN using infrared, X-ray, and rest-frame optical line flux criteria described in Coil et al. (2015), Azadi et al. (2017, 2018), and Leung et al. (2019). The public MOSDEF emission line catalog contains \( 3\sigma \) emission line detections at a limiting line flux of \( \sim 1 \times 10^{17} \text{ erg s}^{-1} \text{ cm}^{-2} \). We distinguish between two samples of MOSDEF galaxies in this work. The first, which we will refer to as MOSDEF Gold, is a sample of 18 galaxies with simultaneous H\( \alpha \) and H\( \beta \) detections used in Section 3 to measure the central tendency and distribution width of the dust attenuation ratio \( Q_{sg} \). The second, which we refer to more generally as the MOSDEF sample, is the set of 263 MOSDEF galaxies with H\( \alpha \) detections. These are used to compare against the 3D-HST and FMOS-COSMOS samples using an assumed \( Q_{sg} = 0.43 \) based on the results detailed in Section 3.

All three spectroscopic surveys are supplemented with Gaussian Process SED fits (Iyer et al. 2019) to photometry from CANDELS fields. This flexible, non-parametric SFH reconstruction method produces star formation rates, stellar masses, and lookback times corresponding to quantiles of the galaxy’s observed stellar mass. Star formation histories are reconstructed using a brute-force Bayesian approach with a large pregrid of model SEDs, producing smooth star formation histories that are not reliant on any particular functional form. This method also produces measurements of the dust attenuation and other standard galaxy physical properties. We apply a \( S/N > 1 \) cut for the H\( \alpha \) and SED measured SFRs for all surveys to exclude noise-dominated objects.

2.4. Simulations

We utilize two large-volume cosmological simulations to compare directly against observations: publicly available runs for MUFASA (Dave et al. 2017) and the Santa Cruz Semi-Analytic Model (SAM; Somerville et al. 2008, 2015; Yung et al. 2019; Brennan et al. 2017). The MUFASA hydrodynamical simulation attempts to directly model the fluid dynamics of the interstellar medium (ISM), intergalactic medium (IGM), and intercluster medium (ICM) in and around galaxies. The Santa Cruz SAM instead achieves greater computational efficiency by applying physically motivated recipes to galaxies in dark matter halos computed from N-body simulations to determine their physical properties (Somerville et al. 2015). Star formation histories are calculated using the Schmidt-Kennicutt law (Kennicutt et al. 1998; Kennicutt 1989) for normal quiescent star formation in isolated disks while merger-driven bursty star formation is based on recipes derived from hydrodynamical simulations of galaxy mergers in Robertson et al. (2006) and Hopkins et al. (2009).

2.5. Generating Tracer Star Formation Rates

The process by which we generate tracer star formation rates is described in detail in Broussard et al. (2019), however we include a brief summary here. Calculations of the detailed time-response of various star formation tracers such as the H\( \alpha \) emission line and NUV flux density is made possible with the use of Flexible Stellar Population Synthesis (FSPS; Conroy et al. 2009; Conroy & Gunn 2010), a software package designed to generate realistic spectra of stellar populations from an input simulated SFH and physical properties such as dust attenuation, stellar initial mass function (IMF), and metallicity. FSPS combines calculations of stellar evolution with stellar spectral libraries to produce spectra for simple stellar populations. FSPS implements CLOUDY (Byler et al. 2017) for calculating nebular
Figure 2. Balmer decrement \( (F_{\text{H}\alpha,\text{uncorr}}/F_{\text{H}\beta,\text{uncorr}}) \) and corresponding \( Q_{\text{sg}} \) (assuming \( E(B-V)_{\text{stars}} = 0.1 \)) vs. \( \text{H}\beta \) SFR signal-to-noise ratio for MOSDEF galaxies with simultaneous \( \text{H}\alpha \) and \( \text{H}\beta \) detections. The red horizontal line indicates a Balmer decrement of 2.86, which is expected in the absence of dust. The green shaded region indicates \( \text{SFR}_{\text{H}\beta} \text{ SNR} > 5 \), which we use to select the MOSDEF Gold sample used to measure the distribution of \( Q_{\text{sg}} \).

Emission. CLOUDY is a photo-ionization code that assumes a constant density spherical shell of gas surrounding the stellar population. Throughout this work, we use the input parameters of FSPS to specify a Chabrier IMF (Chabrier 2003), include nebular line and continuum emission, implement a Calzetti dust law (Calzetti et al. 2000), and assume a gas-phase metallicity of \( Z = 0.2Z_\odot \).

3. MEASURING VARIATIONS IN \( Q_{\text{sg}} \)

The presence of two hydrogen lines for 111 galaxies in the MOSDEF survey enables direct dust correction of the \( \text{H}\alpha \) flux via the Balmer decrement under the assumption of Case B recombination. Combined with UV dust attenuation measurements from SED fitting, we have the means to constrain the detailed distribution of \( Q_{\text{sg}} = E(B-V)_{\text{stars}}/E(B-V)_{\text{gas}} \) values that represent the ratio of dust reddening between stellar and nebular light.
Figure 3. Plots showing the evolution of the Santa Cruz SAM (top row) and Mufasa (bottom row) mock catalogs based on 3D-HST observations. The detailed descriptions of each step involved in creating the mock catalog are described in Section 4.

Solving for dust reddening under the assumption of Case B recombination yields the equation for converting from Balmer decrement to \( E(B - V)_{\text{gas}} \):

\[
E(B - V)_{\text{gas}} = \log_{10} \left( \frac{F_{\text{H}\alpha}}{2.86F_{\text{H}\beta}} \right) \times 2.5(k_{\text{H}\beta} - k_{\text{H}\alpha})^{-1}
\] (1)

where \( F_{\text{H}\alpha} \) and \( F_{\text{H}\beta} \) are the uncorrected emission line fluxes, and \( k_{\text{H}\alpha} \) and \( k_{\text{H}\beta} \) are the values of the assumed dust attenuation curve at the \( \text{H}\alpha \) (\( \lambda = 6563\text{Å} \)) and \( \text{H}\beta \) (\( \lambda = 4863\text{Å} \)) wavelengths respectively.

In practice, we find that 31 of the starting 111 galaxies exhibit Balmer ratios in which \( F_{\text{H}\alpha}/F_{\text{H}\beta} < 2.86 \), as we show in Figure 2. Similarly puzzling Balmer ratios are discussed in Boogaard et al. (2018) as being found in other spectroscopic surveys and have several possible explanations, but in our case, these objects typically have high relative uncertainty and are likely the result of scatter. As a result, we limit the sample used to calculate the \( Q_{sg} \) distribution to only those galaxies with \( \text{SFR}_{\text{H}\beta}/\sigma_{\text{SFR}_{\text{H}\beta}} > 5 \), eliminating all but two objects with such Balmer decrements, and leaving 18 galaxies in total (the MOSDEF Gold sample). From these 18 galaxies, we find \( \text{med}(Q_{sg}) = 0.43 \), consistent with dust measurements in low-redshift star forming galaxies (Calzetti et al. 2000) and the width of the \( Q_{sg} \) distribution \( \sigma_{Q_{sg}} = 0.33 \), calculated via a robust estimator.

Here, it is also important to note that because \( Q_{sg} = E(B - V)_{\text{stars}}/E(B - V)_{\text{gas}} \), Balmer decrements greater than 2.86 yield \( E(B - V)_{\text{gas}} > 0 \) (and therefore \( Q_{sg} > 0 \)) while Balmer decrements less than 2.86 yield \( E(B - V)_{\text{gas}} < 0 \) (and therefore \( Q_{sg} < 0 \)). Further, a Balmer decrement of \( \sim 2.86 \) indicates a galaxy with nearly zero nebular dust reddening, meaning that \( Q_{sg} \) tends toward negative infinity as the Balmer decrement tends toward 2.86 from below, while \( Q_{sg} \) tends toward positive infinity as the Balmer decrement tends toward 2.86 from above, as is indicated by the right axis in Figure 2.

4. MOCK CATALOG GENERATION

In order to better understand the predictions of the SAM and Mufasa simulations, we follow Broussard et al. (2019) by creating a mock catalog from each simulation that is designed to mimic the observational systematics of each observed data set. Throughout the mock catalog generation process, each time bin of a simulated galaxy’s SFR is treated individually, meaning that a single galaxy’s star formation history can contribute multiple “independent” SFR values to the mock catalog, and for the purpose of the mock catalog, each of these is treated as a separate galaxy (though we will later restrict the sample to be within \( |\Delta z| < 0.2 \) of each survey’s mean redshift). While a description of each step in the mock catalog generation process for each of the observed samples is detailed below, a visual summary is shown in Figure 3 for 3D-HST, Figure 4 for FMOS-COSMOS, and Figure 5 for the MOSDEF sample, assuming \( Q_{sg} = 0.43 \) for all three. These summary figures demonstrate the progression in SFR_{NUV} vs. SFR_{H\alpha} as each selection effect is modeled. For clarity, throughout the remainder of this work we will refer to the addition of the effects of dust to simulated fluxes as dust attenuation while the removal of the effects of dust from simulated or observed fluxes will be referred to as dust correction.
Table 1. Schechter Function Parameters

| Sample            | Avg. Redshift | SFR_{*,-1.6} (M_☉ yr^{-1}) | log_{10}(Φ_{*,-1.6}) (Mpc^{-3}) | α     |
|------------------|---------------|-----------------------------|----------------------------------|-------|
| 3D-HST           | 1.00          | 13.98                       | -2.67                            | -1.65 |
| FMOS-COSMOS      | 1.59          | 30.89                       | -2.75                            | -1.61 |
| MOSDEF (1.37 < z < 1.7) | 1.54          | 28.23                       | -2.73                            | -1.59 |

Note—A table of the average redshifts for each survey as well as the interpolated Schechter parameters used to select the initial sample of mock galaxies based on their SFR_{Hα} described in Section 4.1.

Figure 4. Plots showing the evolution of the Santa Cruz SAM (top row) and MUFASA (bottom row) mock catalogs based on FMOS-COSMOS observations. The detailed descriptions of each step involved in creating the mock catalog are described in Section 4.

4.1. Initial SFR_{Hα} Draws

To form our initial sample of galaxies for the mock catalog, we begin by drawing SFR measurements such that our starting galaxy sample matches the observed star formation rate function (SFRF) from the High Redshift (Z) Emission Line Survey (HIZELS; Geach et al. 2008; Sobral et al. 2012, 2013). HIZELS is a narrowband survey and consequently, star formation rate functions are available as Schechter function fits at z = 0.4, 0.84, 1.47, and 2.23. Because the 3D-HST sample has an average redshift of z ≈ 1, we interpolate the SFRF parameters between z = 0.84 and z = 1.47 to approximate the SFRF at z ∼ 1. We perform a similar interpolation for FMOS-COSMOS using the average redshift of z ∼ 1.59 and for MOSDEF using z ∼ 1.54. The resulting Schechter function parameters are summarized in Table 1.

We draw galaxies from the SAM and MUFASA simulations by normalizing the SFRF to 1 by integrating over 10^{-3} < SFR/SFR_{*} < 10^3, effectively producing a probability distribution function (PDF) such that p(SFR) ∝ φ(SFR). When drawing SFRs from this distribution, we select an SFR from the SAM or MUFASA simulation that lies within 0.5 dex of the drawn SFR and with |z_{sim} - z_{obs}| < 0.2 where z_{sim} is the redshift of the matched SAM or MUFASA SFR and z_{obs} is the average redshift of the particular mock’s corresponding survey. Finally, SFR_{Hα}, SFR_{NUV}, and M_{*} for the matched galaxy are all scaled such that SFR_{Hα} exactly matches the original SFRF-drawn SFR. This process is repeated for 10^5 galaxy draws from this distribution for each of the mock catalogs, producing the initial distribution of galaxies shown in the leftmost panels of Figures 3, 4, and 5.

4.2. Dust Attenuation

We use FSPS to derive the relationship between star formation rate and the observed Hα luminosity (L_{Hα}) and NUV luminosity density (L_{ν,NUV}) by generating dust-free galaxy spectra for a constant star formation rate, finding that ν_{NUV}L_{ν,NUV} = SFR_{NUV} × 10^{42.94} erg s^{-1} Hz^{-1}/(M_☉ yr^{-1}) and L_{Hα} = SFR_{Hα} × 10^{41.20} erg s^{-1}/(M_☉ yr^{-1}).
these relationships, we convert the mock $\text{SFR}_{\text{H}\alpha}$ and $\text{SFR}_{\text{NUV}}$ into their corresponding flux and flux density respectively so that dust attenuation can be applied.

For each observed sample, we perform fits of various galaxy properties against $A_V$ to determine the most robust relationship. For 3D-HST, we find a tight correlation between $\text{SFR}_{\text{H}\alpha}$ and $A_V$. For FMOS-COSMOS, we find the strongest correlation to be between $A_V$ and $\text{SFR}_{\text{NUV}}$. Meanwhile, for MOSDEF, we find the strongest correlation to be between $A_V$ and $M_\ast$. For each of these cases, we fit a Gaussian mixture model (GMM) to each relationship, enabling random assignment of $A_V$ values to mock galaxies in a way that preserves the covariance between $A_V$ and the best-correlated galaxy property.

Because of the addition of the MOSDEF sample to this analysis that includes simultaneous detections of the $\text{H}\alpha$ and $\text{H}\beta$ emission lines, we are able to characterize the distribution of $Q_{\text{sg}}$ values as shown in Figure 2. We assign values of $Q_{\text{sg}}$ (which we call $Q_{\text{sg, intrinsic}}$) to drawn galaxies using a Gaussian distribution with a $\mu = 0.43$ and $\sigma = 0.33$. We set a minimum assignable value of 0.05 to avoid negative values of $Q_{\text{sg, intrinsic}}$, which would imply an unphysical, negative $E(B-V)_{\text{gas}}$.

### 4.3. Flux Uncertainties

We approximate the uncertainty of a given simulated galaxy’s $\text{H}\alpha$ flux by first fitting the relationship between the $\text{H}\alpha$ flux its associated uncertainty for each of the three observed samples. This process is described in detail in Broussard et al. (2019), and we perform a similar fit here to each of the 3D-HST, FMOS-COSMOS, and MOSDEF samples to assign realistic flux uncertainties to mock galaxy $\text{H}\alpha$ fluxes and NUV flux densities.

#### 4.4. Flux Limits

In order to model the various observational limits and selection effects of each survey, we use FSPS to directly estimate the flux density in detection bands using the detailed star formation history and assigned $A_V$ value for each mock galaxy. The detailed selection criteria are listed below for each survey.

The 3D-HST survey requires a detection in the F140W filter, which is used as the calibration image for the grism data. This corresponds to a 25.8 magnitude limit in the F140W band. We further exclude all mock galaxies that do not meet the $\text{SNR} > 8$ grism detection limit of $8.8 \times 10^{-17} \text{erg s}^{-1} \text{cm}^{-2}$.

The FMOS-COSMOS sample uses a series of sample selections. A cut in $(B - z) - (z - K)$ selects star forming galaxies based on the findings of Daddi et al. (2004). A cut in K-band magnitude of $K_s < 23$ is based on the depth and completeness of CFHT WIRCam observations (McCracken et al. 2010). Targets are also excluded if the estimated $\text{H}\alpha$ flux is below the $3\sigma$ detection limit of $4 \times 10^{-17} \text{erg s}^{-1} \text{cm}^{-2}$, with the flux estimated by treating the UV SFR as an $\text{H}\alpha$ SFR and combining it with the spectroscopic redshift to yield an observed line flux. Additionally, a mass cut of $\log_{10}(M_\ast) > 9.8$ is implemented to our mock galaxies to match the targeted mass range of the survey.

MOSDEF implements a magnitude cut of $H < 24$ for the low-redshift interval $1.37 < z < 1.70$ when selecting targets. Although the initial pool of targets is itself composed of objects with detections in the 3D-HST photometric
catalog (Skelton et al. 2014), we do not implement the combined $F125W + F140W + F160W$ image detection criteria as Skelton et al. (2014) notes that the sample is $\sim 90\%$ complete even in the shallow CANDELS fields when selecting only objects with $H < 25$. We also exclude mock objects that are not sufficiently bright to meet the $SNR > 3$ MOSDEF detection limit of $\sim 1.7 \times 10^{-17}$ erg s$^{-1}$ cm$^{-2}$.

4.5. Dust Correction

With the observed systematics implemented, we correct for dust in each mock galaxy. Each mock catalog corrects for dust using a “measured” value of $A_V$ that has scatter added based on the relationship between $A_V$ and $\sigma_{A_V}$, assuming $Q_{sg,measured} = 0.43$ across all simulated observations. As a result, the effect of adding dust attenuation to the mocks and correcting it in this later step for galaxies that are not excluded by survey selection effects is an increase in flux scatter relative to the intrinsic values due to the associated $A_V$ uncertainty and measurement errors that arise from assuming a single value of $Q_{sg,assumed} = 0.43$. For reference $Q_{sg} > Q_{sg,assumed}$ will yield $F_{H\alpha,meas} > F_{H\alpha}$ for the same $A_V$ (and vice-versa).

4.6. Matching SFR Measurement Uncertainties

To ensure that the mock catalogs are generating systematic scatter that is comparable to that of each corresponding observed sample, we track each galaxy through the mock catalog generation process. After applying each of the previous steps, we compute the difference between each remaining galaxy’s resulting $SFR_{H\alpha}$ ($SFR_{NUV}$) and the $SFR_{H\alpha}$ ($SFR_{NUV}$) it was assigned during the initial draw in Section 4.1, producing an effective mock observational error. We then bin galaxies by $SFR_{H\alpha}$ and $SFR_{NUV}$, computing the average uncertainty of the observed galaxy in each bin and the standard deviation of the mock observational error and add Gaussian scatter to the relevant mock SFRs with $\sigma_{SFR,added} = \sqrt{\sigma_{SFR,obs}^2 - \sigma_{SFR, mock}^2}$. This causes the mock observational error in each bin to rise to match the average measurement uncertainty, thus ensuring that the mocks do not underestimate the scatter inherent in computing SFRs from observations. In practice, this adds scatter to fewer than half of the bins, and the added scatter well under 1 dex.

5. CORRECTING FOR SYSTEMATIC SCATTER IN $\eta$

Thus far, we have endeavored to apply forward modeling techniques to simulated data in an effort to compare burstiness between the mock and observed samples. Here, we will show that, with accurate knowledge of observational uncertainties, it is also possible to estimate the underlying burstiness of an observed sample, which can then be compared directly against the $\eta$ distribution for simulated galaxies (e.g., after making a reasonable cut in $M_\star$).

For a single galaxy, we can express the observed burst indicator value $\eta_{obs}$ as the sum of some true underlying burst indicator value $\eta$ and the measurement error $\delta_{\eta}$:

$$\eta_{obs} = \eta + \delta_{\eta}.$$ (2)
### Table 2. \( \eta \) Distribution Statistics

| Survey             | Sample          | \( \bar{\eta} \) | \text{NMAD}(\eta) |
|--------------------|-----------------|-------------------|-------------------|
| 3D-HST             | Observed        | 0.175             | 0.208             |
|                    | SAM mock        | 0.165             | 0.275             |
|                    | MUFASA mock     | 0.193             | 0.295             |
| FMOS-COSMOS        | Observed        | 0.302             | 0.246             |
|                    | SAM mock        | 0.116             | 0.198             |
|                    | MUFASA mock     | 0.146             | 0.179             |
| MOSDEF (1.37 < \( z \) < 1.7) | Observed       | 0.054             | 0.259             |
|                    | SAM mock        | -0.070            | 0.235             |
|                    | MUFASA mock     | 0.004             | 0.151             |

Note— Descriptive statistics of the various \( \eta \) distributions shown in Figure 6. Here, we denote the median of \( \eta \) as \( \bar{\eta} \).

Shifting now to a view of a population of such galaxies, we can describe the variance in \( \eta_{\text{obs}} \) in terms of the intrinsic \( \eta \) variance and the “explainable” portion of the observed variance for an ensemble of galaxies using Equation 2:

\[
\text{Var}(\eta_{\text{obs}}) = \text{Var}_{\text{int}}(\eta) + \text{Var}(\delta \eta) = \text{Var}_{\text{int}}(\eta) + \text{Var}_{\text{exp}}(\eta_{\text{obs}}). \tag{3}
\]

#### 5.1. Dominant SFR Uncertainties

In the case where SFR uncertainties dominate the \( \eta \) measurement uncertainty, it is possible to break \( \eta \) into its constituent parts using its definition as the logarithm of the ratio of two SFRs:

\[
\eta_{\text{obs}} = \log_{10} \left( \frac{\text{SFR}_{\text{H}\alpha,\text{obs}}}{\text{SFR}_{\text{NUV,obs}}} \right). \tag{4}
\]

This yields the following equation for the explained variance in \( \eta_{\text{obs}} \) in terms of the SFR measurement errors:

\[
\text{Var}_{\text{exp}}(\eta_{\text{obs}}) = \text{Var}(\delta \log_{10} \text{SFR}_{\text{H}\alpha,\text{obs}}) + \text{Var}(\delta \log_{10} \text{SFR}_{\text{NUV,obs}}) - 2 \text{Cov}(\delta \log_{10} \text{SFR}_{\text{H}\alpha,\text{obs}}, \delta \log_{10} \text{SFR}_{\text{NUV,obs}}). \tag{5}
\]

Here, we set the final term to zero because the measurement errors for each star formation rate should be linearly independent, and we can rewrite the expression in terms of the galaxies’ SFR uncertainties:

\[
\text{Var}_{\text{exp}}(\eta_{\text{obs}}) = \text{Avg}(\sigma^2_{\log_{10} \text{SFR}_{\text{H}\alpha,\text{obs}}}) + \text{Avg}(\sigma^2_{\log_{10} \text{SFR}_{\text{NUV,obs}}}). \tag{6}
\]

While this method turns out to be insufficient for describing the variance in \( \eta \) because it does not take into account detailed dust systematics, it is a useful starting point for the more complex derivations to come.

#### 5.2. Flux and Dust Attenuation Uncertainties with Known Balmer Decrement

We consider two differing regimes - galaxy samples with and without measured Balmer decrements - when calculating the scatter in \( \eta_{\text{obs}} \). Starting with the case of known Balmer decrements, we can expand the two star formation rates into quantities that are either directly observed or inferred through redshift fitting and spectral energy distribution (SED) fitting yields the following sequence of equations, culminating in the redefinition of \( \eta \) below in terms of the galaxy’s observed flux (density), the SFR conversion factors, the chosen dust law, and the dust attenuation of light.
emitted by nebular and stellar light respectively.

\[
\begin{align*}
\text{SFR}_{\text{H}\alpha} &= C_{\text{H}\alpha} (4\pi d_{\text{obs}}^2) F_{\text{H}\alpha} \times 10^{0.4 k_{\text{H}\alpha} E(B-V)_{\text{gas}}} \\
\text{SFR}_{\text{NUV}} &= C_{\text{NUV}} (4\pi d_{\text{obs}}^2) \nu_{\text{F},\text{NUV}} \times 10^{0.4 k_{\text{NUV}} E(B-V)_{\text{stars}}} \\
\frac{\text{SFR}_{\text{H}\alpha}}{\text{SFR}_{\text{NUV}}} &= \frac{C_{\text{H}\alpha} F_{\text{H}\alpha}}{C_{\text{NUV}} \nu_{\text{F},\text{NUV}}} \times 10^{0.4 k_{\text{H}\alpha} E(B-V)_{\text{gas}} - k_{\text{NUV}} E(B-V)_{\text{stars}}} \\
\eta &= \log_{10} \left( \frac{C_{\text{H}\alpha} F_{\text{H}\alpha}}{C_{\text{NUV}} \nu_{\text{F},\text{NUV}}} \right) \\
&+ 0.4 \left[ k_{\text{H}\alpha} E(B-V)_{\text{gas}} - k_{\text{NUV}} E(B-V)_{\text{stars}} \right] \\
&= \log_{10} \left( \frac{C_{\text{H}\alpha} F_{\text{H}\alpha}}{C_{\text{NUV}} \nu_{\text{F},\text{NUV}}} \right) \\
&+ 0.4 k_{\text{H}\alpha} E(B-V)_{\text{gas}} + 0.4 k_{\text{NUV}} E(B-V)_{\text{stars}}.
\end{align*}
\]

Under the assumption of a Calzetti et al. (2000) dust law \((k_{\text{H}\alpha}, k_{\text{NUV}})\) and Kennicutt & Evans (2012) SFR conversion factors \((C_{\text{H}\alpha}, C_{\text{NUV}})\), we are left with four measured parameters with associated uncertainties: \(\log_{10} F_{\text{H}\alpha}, \log_{10} F_{\text{NUV}}, E(B-V)_{\text{gas}},\) and \(E(B-V)_{\text{stars}},\) where we have adopted \(F_{\text{NUV}} \equiv \nu F_{\text{nu,NUV}}\) to simplify the subsequent notation. It is apparent by examining Equations 2 and 4 that we can replace each of these four parameters with a term of the form \(x + \delta_x\), representing the sum of its intrinsic value and the individual observational error. This resulting equation would then describe \(\eta_{\text{obs}}\) for the hypothetical galaxy, with the observational error terms summing to \(\delta_{\eta}\).

Returning to Equation 10 we can express \(\eta_{\text{obs}}\) for a single galaxy in terms of true values and their observational errors:

\[
\eta_{\text{obs}} = \log_{10}(C_{\text{H}\alpha}) + \log_{10}(F_{\text{H}\alpha}) + \delta_{\log_{10} F_{\text{H}\alpha}} + 0.4 k_{\text{H}\alpha} E(B-V)_{\text{gas}} + 0.4 k_{\text{NUV}} E(B-V)_{\text{stars}} - \log_{10}(C_{\text{NUV}} \nu) - \log_{10}(F_{\text{NUV}}) - \log_{10}(F_{\text{NUV}}) - \log_{10}(F_{\text{NUV}}) \\
- 0.4 k_{\text{H}\alpha} E(B-V)_{\text{gas}} + 0.4 k_{\text{NUV}} E(B-V)_{\text{stars}} \right). \tag{11}
\]

We then continue by subtracting away the true burst indicator \(\eta\) and taking the variance to isolate the explained variance from observations \(\text{Var}_{\text{exp}}(\eta_{\text{obs}})\) (i.e., the scatter in \(\eta_{\text{obs}}\) caused by observational uncertainties), which contains four variance terms and six covariance terms:

\[
\begin{align*}
\text{Var}_{\text{exp}}(\eta_{\text{obs}}) &= \text{Var}(\log_{10} F_{\text{H}\alpha}) + \text{Var}(\delta_{\log_{10} F_{\text{H}\alpha}}) + (0.4 k_{\text{H}\alpha})^2 \text{Var}(\delta_{E(B-V)_{\text{gas}}}) \\
&+ (0.4 k_{\text{NUV}})^2 \text{Var}(\delta_{E(B-V)_{\text{stars}}}) + 2 \text{Cov}(\delta_{\log_{10} F_{\text{H}\alpha}}, \delta_{\log_{10} F_{\text{NUV}}}) \\
&+ 2(0.4 k_{\text{H}\alpha}) \text{Cov}(\delta_{\log_{10} F_{\text{H}\alpha}}, \delta_{E(B-V)_{\text{gas}}}) \\
&+ 2(0.4 k_{\text{NUV}}) \text{Cov}(\delta_{\log_{10} F_{\text{NUV}}}, \delta_{E(B-V)_{\text{gas}}}) \\
&+ 2(0.4 k_{\text{H}\alpha}) \text{Cov}(\delta_{\log_{10} F_{\text{NUV}}}, \delta_{E(B-V)_{\text{stars}}}) \\
&+ 2(0.4 k_{\text{NUV}}) \text{Cov}(\delta_{E(B-V)_{\text{gas}}}, \delta_{E(B-V)_{\text{stars}}}) \\
&+ 2(0.16 k_{\text{H}\alpha} k_{\text{NUV}}) \text{Cov}(\delta_{E(B-V)_{\text{stars}}}, \delta_{E(B-V)_{\text{gas}}}). \tag{12}
\end{align*}
\]

The observational errors associated with measuring \(F_{\text{H}\alpha}\) and \(F_{\text{NUV}}\) are uncorrelated, as are the errors associated with each of the \(E(B-V)\) terms. While there is possibly some covariance between \(\delta_{\log_{10} F_{\text{NUV}}}\) and \(\delta_{E(B-V)_{\text{stars}}}\) because of the dependence of dust measurements on UV emission, dust corrections derived using SED fitting (as is the case for each of our samples) are not strongly dependent on any single band. As a result, we set each of these covariance terms
to zero, leaving only the four variance terms:

\[
\text{Var}_{\text{exp}}(\eta_{\text{obs}}) = \text{Var}(\delta_{\log_{10} F_{\text{H}}}) \\
+ \text{Var}(\delta_{\log_{10} F_{\text{NUV}}}) \\
+ (0.4k_{\text{H}})^2 \text{Var}(\delta_{E(B-V)_{\text{gas}}}) \\
+ (0.4k_{\text{NUV}})^2 \text{Var}(\delta_{E(B-V)_{\text{stars}}}).
\]  

(13)

Finally, we can express the above equation in terms of the reported uncertainties to arrive at an expression for the scatter added to \( \eta \) as a result of observational uncertainties:

\[
\text{Var}_{\text{exp}}(\eta_{\text{obs}}) = \text{Avg}(\sigma_{\log_{10} F_{\text{H}}}^2) \\
+ \text{Avg}(\sigma_{\log_{10} F_{\text{NUV}}}^2) \\
+ (0.4k_{\text{H}})^2 \text{Avg}(\sigma_{E(B-V)_{\text{gas}}}^2) \\
+ (0.4k_{\text{NUV}})^2 \text{Avg}(\sigma_{E(B-V)_{\text{stars}}}^2).
\]  

(14)

5.3. Flux and Dust Attenuation Uncertainties without Balmer Decrements

When Balmer Decrements are not available for an ensemble of galaxies, we can instead use the relationship \( E(B-V)_{\text{gas}} = E(B-V)_{\text{stars}}/Q_{\text{sg}} \) to redefine Equation 10 in terms of \( F_{\text{H}} \), \( F_{\text{NUV}} \), \( E(B-V)_{\text{stars}} \), and \( Q_{\text{sg}} \):

\[
\eta = \log_{10} \left( \frac{C_{\text{H}} F_{\text{H}}}{C_{\text{NUV}} \nu F_{\text{NUV}}} \right) \\
+ 0.4E(B-V)_{\text{stars}} \left[ k_{\text{H}} Q_{\text{sg}}^{-1} - k_{\text{NUV}} \right].
\]  

(15)

Continuing, this gives the following relationship for \( \eta_{\text{obs}} \):

\[
\eta_{\text{obs}} = \log_{10} \left( \frac{C_{\text{H}}}{\nu C_{\text{NUV}}} \right) \\
+ \log_{10} F_{\text{H}} + \delta_{\log_{10} F_{\text{H}}} \\
- \log_{10} F_{\text{NUV}} - \delta_{\log_{10} F_{\text{NUV}}} \\
+ 0.4(k_{\text{H}} Q_{\text{sg}}^{-1} - k_{\text{NUV}})E(B-V)_{\text{stars}} \\
+ 0.4(k_{\text{H}} Q_{\text{sg}}^{-1} - k_{\text{NUV}})\delta_{E(B-V)_{\text{stars}}} \\
+ 0.4k_{\text{H}} E(B-V)_{\text{stars}} \delta Q_{\text{sg}}^{-1} \\
+ 0.4k_{\text{H}} \delta_{E(B-V)_{\text{stars}}} \delta Q_{\text{sg}}^{-1}
\]  

(16)
Similar to the case with known Balmer decrements, we subtract away $\eta$ and compute the variance to isolate the explainable variance of $\eta_{\text{obs}}$, finding:

$$
\text{Var}_{\text{exp}}(\eta_{\text{obs}}) = \text{Var}(\delta_{\log_{10} F_{\text{H}\alpha}}) \\
+ \text{Var}(\delta_{\log_{10} F_{\text{NUV}}}) \\
+ 0.4(k_{\text{H}_{\alpha}}Q_{\alpha}^{-1} - k_{\text{NUV}})^2 \text{Var}(\delta E(B-V)_{\text{stars}}) \\
+ 0.4k_{\text{H}_{\alpha}}^2 \text{Var}(E(B-V)_{\text{stars}}) \delta_{Q_{\alpha}^{-1}} \\
+ 0.4k_{\text{H}_{\alpha}}^2 \text{Var}(\delta E(B-V)_{\text{stars}}) \\
- 2 \text{Cov}(\delta_{\log_{10} F_{\text{H}\alpha}}, \delta_{\log_{10} F_{\text{NUV}}}) \\
+ 2[0.4(k_{\text{H}_{\alpha}}Q_{\alpha}^{-1} - k_{\text{NUV}})] \text{Cov}(\delta_{\log_{10} F_{\text{H}\alpha}}, \delta E(B-V)_{\text{stars}}) \\
+ 2[0.4k_{\text{H}_{\alpha}}] \text{Cov}(\delta_{\log_{10} F_{\text{H}\alpha}}, E(B-V)_{\text{stars}}) \delta_{Q_{\alpha}^{-1}} \\
+ 2[0.4k_{\text{H}_{\alpha}}] \text{Cov}(\delta_{\log_{10} F_{\text{H}\alpha}}, \delta E(B-V)_{\text{stars}}) \\
- 2[0.4k_{\text{H}_{\alpha}}] \text{Cov}(\delta_{\log_{10} F_{\text{NUV}}}, \delta_{Q_{\alpha}^{-1}}) \delta E(B-V)_{\text{stars}} \\
+ 2[0.16k_{\text{H}_{\alpha}}(k_{\text{H}_{\alpha}}Q_{\alpha}^{-1} - k_{\text{NUV}})] \text{Cov}(\delta E(B-V)_{\text{stars}}, E(B-V)_{\text{stars}}) \delta_{Q_{\alpha}^{-1}} \\
+ 2[0.16k_{\text{H}_{\alpha}}(k_{\text{H}_{\alpha}}Q_{\alpha}^{-1} - k_{\text{NUV}})] \text{Cov}(\delta E(B-V)_{\text{stars}}, \delta E(B-V)_{\text{stars}}) \delta_{Q_{\alpha}^{-1}} \\
+ 2[0.4k_{\text{H}_{\alpha}}]^2 \text{Cov}(E(B-V)_{\text{stars}}) \delta_{Q_{\alpha}^{-1}}, \delta E(B-V)_{\text{stars}}) \delta_{Q_{\alpha}^{-1}}).
$$

Again, each of the covariance terms without a repeated operand (i.e., Cov($\delta X$, $\delta Y$) or Cov($\delta X$, $\delta Y$ $\delta Z$)) can be set to zero for similar reasons to the previous case. The fifth variance term can be expanded using the relationship $\text{Var}(\delta X$ $\delta Y$) = $E(\delta X)^2 \text{Var}(\delta Y) + E(\delta Y)^2 \text{Var}(\delta X) + \text{Var}(\delta X) \text{Var}(\delta Y)$ under the assumption that $\delta X$ and $\delta Y$ are independent, and all covariance terms with repeated operands will reduce to Cov($\delta X$, $\delta X$ $\delta Y$) = Var($\delta X$) $E[\delta Y]$ and Cov($\delta X$ $\delta Y$, $\delta X$ $\delta Z$) = Var($\delta X$) $E[\delta Y]$ $E[\delta Z]$ under a similar independence assumption between all terms. This assumption is possible because none of the measured values (e.g., $E(B-V)$) are strongly dependent on any single observable (particularly when measured using SED fitting). This gives:

$$
\text{Var}_{\text{exp}}(\eta_{\text{obs}}) = \text{Var}(\delta_{\log_{10} F_{\text{H}\alpha}}) \\
+ \text{Var}(\delta_{\log_{10} F_{\text{NUV}}}) \\
+ 0.4(k_{\text{H}_{\alpha}}Q_{\alpha}^{-1} - k_{\text{NUV}})^2 \text{Var}(\delta E(B-V)_{\text{stars}}) \\
+ 0.4k_{\text{H}_{\alpha}}^2 \text{Var}(E(B-V)_{\text{stars}}) \delta_{Q_{\alpha}^{-1}} \\
+ 0.4k_{\text{H}_{\alpha}}^2 \text{Var}(\delta E(B-V)_{\text{stars}}) \\
- 2[0.16k_{\text{H}_{\alpha}}(k_{\text{H}_{\alpha}}Q_{\alpha}^{-1} - k_{\text{NUV}})] \text{Var}(\delta E(B-V)_{\text{stars}}) \\
+ 2[0.4k_{\text{H}_{\alpha}}]^2 \text{Var}(\delta E(B-V)_{\text{stars}}) \delta_{Q_{\alpha}^{-1}}) \delta E(B-V)_{\text{stars}})
$$

Because the distribution of observational errors in these measured quantities is assumed to be approximately symmetric, all expectation values of $\delta$ terms above can be set to zero, and we can reformulate the equation in terms of the
Table 3. Variance Subtraction Results

| Sample       | Var_{obs}(\eta_{obs}) | Var_{exp}(\eta_{obs}) | Var_{int}(\eta) | \sigma_{int}(\eta) |
|--------------|------------------------|------------------------|-----------------|-------------------|
| 3D-HST       | 0.043                  | 0.034                  | 0.010           | 0.100             |
| FMOS-COSMOS  | 0.058                  | 0.054                  | 0.004           | 0.063             |
| MOSDEF       | 0.073                  | 0.047                  | 0.026           | 0.161             |

Note—A table of the observed variance in the \eta_{obs}, explained variance in \eta_{obs} due to uncertainties and scatter in Q_{sg}, and the corrected intrinsic variance in \eta after subtracting away the explained variance.

reported uncertainties to get an analogous expression to Equation 14 for the case of unknown Balmer decrements:

\[
\text{Var}_{\text{exp}}(\eta_{\text{obs}}) = \text{Avg}(\sigma_{\log_{10} F_{\text{H}\alpha}}^2) + \text{Avg}(\sigma_{\log_{10} F_{\text{NUV}}}^2) + [0.4(k_{\text{H}\alpha}Q_{sg}^{-1} - k_{\text{NUV}})]^2 \text{Avg}(\sigma_{E(B-V)\text{stars}}^2) + [0.4k_{\text{H}\alpha}]^2 \text{Avg}(\sigma_{Q_{sg}^{-1}}^2) \text{Avg}(\sigma_{E(B-V)\text{stars}}^2). \tag{19}
\]

Here, \sigma_{Q_{sg}^{-1}}^2 represents the variance of Q_{sg}^{-1} as measured from the MOSDEF Gold sample, which we use to approximate the additional measurement uncertainty introduced by the assumption that all galaxies have Q_{sg} = 0.43.

In summary, Equation 14 describes the relationship between the explainable scatter in \eta_{obs} and observed uncertainties in the case of independent measurements of nebular and stellar dust attenuation. Equation 19, on the other hand, describes the relationship for a sample with no independent measurements of stellar and nebular dust, but a known Q_{sg} distribution. Correspondingly, we will apply Equation 19 to all three samples using the Q_{sg} distribution measured from the MOSDEF Gold sample.

6. RESULTS

6.1. Mock Catalog Approach

With the completion of the mock catalog, we can now compare the observed distribution of the burst indicator \eta against that of the mock catalogs, which we show in Figure 6. Here, we use the Normalized Median Absolute Deviation (NMAD) to estimate the \eta scatter in each panel because it is robust to outliers. We find that small, but significant differences can be seen when comparing the observed histograms against those of the mock catalogs for each sample. We find an over-estimation of the observed burstiness for the 3D-HST sample, but an under-estimation of the burstiness for the FMOS-COSMOS and MOSDEF samples. Further, while the median of \eta is in good agreement with observations for the 3D-HST mock catalog, FMOS-COSMOS and MOSDEF show an under-estimation in the mock catalogs. We summarize the statistics of each \eta distribution in Table 2.

6.2. Variance Estimation Approach

We apply Equation 19 to each of our three observed samples while assuming a constant value of Q_{sg} = 0.43. We expect that each of these samples should produce differing initial estimates for the total scatter in \eta while producing similar estimates of the intrinsic scatter. This process yields intrinsic burstiness measurements of 0.100 dex for 3D-HST, 0.063 dex for FMOS-COSMOS, and 0.161 dex for MOSDEF. A full accounting of the observed scatter, explained scatter, and corrected intrinsic scatter in \eta is given in Table 3.

7. DISCUSSION

7.1. Mock Catalog Offsets

We find good agreement between the shape of observed and mock \eta histograms for the 3D-HST sample; however, we note a discrepancy between distribution medians for MOSDEF and FMOS-COSMOS observations and mock samples.
(see Table 2 for detailed statistics). While this discrepancy could be relieved by assuming \( Q_{sg} \approx 1 \) (see Broussard et al. 2019 for illustrations), this would contradict the measurements of \( Q_{sg} \) carried out in Section 3. One possible explanation of this behavior could be that high-SNR MOSDEF galaxies do not have \( Q_{sg} \) values that are representative of more typical galaxies in MOSDEF and FMOS-COSMOS. It is also possible that increased precision is needed when measuring \( E(B - V)_{stars} \) and \( E(B - V)_{gas} \) to characterize the \( Q_{sg} \) distribution.

### 7.2. Ionizing Photon Production Efficiency

Another potential explanation for this phenomenon could arise from the ionizing photon production efficiency (\( \xi \)). Also known as the Lyman Continuum (LyC) production efficiency represented by \( \xi_{ion} \), the ionizing photon production efficiency is defined as the ratio of the production rate of ionizing photons (denoted as \( N(H^0) \)) to the UV continuum luminosity density \( L_{\nu,\text{UV}} \):

\[
\xi_{ion} = \frac{N(H^0)}{L_{\nu,\text{NUV}}}.
\]  

Shivaei et al. (2018) calculates \( N(H^0) \) by assuming Case B recombination, effectively converting the entire \( H_\alpha \) luminosity into ionizing photons. They use the relation of Leitherer et al. (1995) to perform this conversion:

\[
N(H^0) = \frac{10^{12} L_{H_\alpha}}{1.36}.
\]  

This indicates that \( \xi_{ion} \) can in fact be expressed as a ratio of the \( H_\alpha \) luminosity to the UV continuum luminosity density multiplied by a constant. Thus, the relationship between \( \xi_{ion} \) and \( \eta \) is simply:

\[
\eta = \log_{10}(\xi_{ion}) + \log_{10}(C)
\]

\[
C = \frac{C_{H_\alpha}}{C_{NUV}} \times \frac{1.36}{10^{12}}
\]

As a result, measurements of \( \eta \) are subject to the same systematic effects that should be accounted for when measuring \( \xi_{ion} \). Shivaei et al. (2018) note four primary sources of uncertainty in measuring \( \xi_{ion} \), including dust attenuation within HII regions, the escape fraction of LyC photons, and the dust corrections applied to the observed \( H_\alpha \) luminosities and UV luminosity densities respectively. Dust attenuation within HII regions has the potential to confound both analyses, but is assumed to be a negligible effect because of the strong agreement between \( H_\alpha \) and UV SFRs. Although the LyC escape fraction is expected to be \( f_{\text{esc}} \ll 1 \), a non-zero LyC escape fraction could also bias both analyses, and is unaccounted for in this work. Finally, the dust attenuation of \( H_\alpha \) and UV luminosities is accounted in a similar manner by Shivaei et al. (2018) who use the Balmer decrement to correct for nebular dust attenuation and SED fitting to correct for UV dust attenuation. The ionizing photon efficiency is also dependent on recent star formation and, while an analysis measuring \( \xi_{ion} \) would typically average over these variations as a systematic effect, we aim to measure them because of their close relationship to burstiness.

### 7.3. Variance Estimation Approach

When we study the contributions of each of the terms in Equation 19 in practice, the fourth term dominates the correction factor across all three samples, causing the magnitude of the correction factor to strongly depend upon each sample’s \( E(B - V) \) values. This causes the FMOS-COSMOS sample’s correction factor to rise above those of the other two samples, as it contains galaxies that are, on average, dustier than the others.

While Table 3 shows a wide range in measurements of the intrinsic variance in \( \eta \), this is largely due to large uncertainties in measurements of Balmer decrements and \( E(B - V)_{stars} \) (and therefore also \( Q_{sg} \)). We estimate the range of the intrinsic burstiness to be \( 0.06 - 0.16 \) dex based on the results of our three observed samples. Similar to our discussion of the mock catalog results, these measurements will likely benefit from a larger sample of galaxies as well as increased precision for measurements of the \( Q_{sg} \) distribution. Early \textit{JWST} surveys will greatly benefit similar analyses in the future, as the Cosmic Evolution Early Release Science (CEERS) program alone, for example, promises deep spectra of several hundred galaxies at 5\( \sigma \) emission line limiting fluxes of \( 1 - 2 \times 10^{-18} \text{ erg s}^{-1} \text{ cm}^{-2} \).
8. CONCLUSIONS

This work adds spectroscopy from FMOS-COSMOS and MOSDEF to the 3D-HST dataset analyzed in Broussard et al. (2019). Most importantly, the addition of MOSDEF carries with it galaxies with simultaneous Hα and Hβ detections, enabling Balmer decrement dust corrections of nebular emission lines. Combined with updated SED fits using the Dense Basis (Iyer et al. 2019) method, we are able to measure the distribution of $Q_{sg}$, finding an average value of $Q_{sg} = 0.43$ and $\sigma_{Q_{sg}} = 0.33$.

We generate updated mock galaxy catalogs using the Santa Cruz SAM and Mufasa cosmological hydrodynamical simulations based on the individual observational constraints for the 3D-HST and FMOS-COSMOS surveys, as well as the MOSDEF sample (those galaxies with a SFR$_{H\beta}$ SNR $> 5$ for which we assume $Q_{sg} = 0.43$, similar to the other two samples). We find decent agreement between the $\eta$ distributions of the observed samples and those of their respective mock catalogs; however, the FMOS-COSMOS and MOSDEF mocks exhibit a shift toward lower $\eta$ and decreased scatter in $\eta$. These differences could have a number of causes, including differences in the ionizing photon production efficiency $\xi_{ion}$, insufficiently precise measurements of the distribution of $Q_{sg}$, and of course genuine differences between simulations and observations that have not previously been probed. $Q_{sg}$ measurement precision could be improved by additional spectroscopy targeting Balmer decrements as well as more precise measurements of the stellar dust attenuation.

We also describe the close relationship between $\eta$ and the logarithm of the ionizing photon efficiency $\xi_{ion}$, which are related by a constant. Most of the systematics that must be accounted for in an analysis studying $\xi_{ion}$ are also present when measuring $\eta$; meanwhile, we note that variations in recent star formation that are a systematic for $\xi_{ion}$ measurements are the target of study of this work.

Beyond the mock catalog approach first introduced in Broussard et al. (2019), this work adds a novel analysis in which the intrinsic variance in $\eta$ can be estimated by accounting for the amount of scatter in $\eta_{obs}$ that is the result of observational uncertainties and variations in $Q_{sg}$. We are able to perform the first measurement of intrinsic scatter in $\eta$, finding an intrinsic scatter of $0.06 - 0.16$ dex. We conclude that more precise measurements of $Q_{sg}$ via stellar dust attenuation or Balmer decrements are likely necessary to improve the precision of this result. We note early JWST surveys will provide more numerous and more precise measurement uncertainties for Hα and Hβ that will greatly benefit similar analyses in the near future.

ACKNOWLEDGEMENTS

AB and EG acknowledge support from NASA ADAP grant 80NSSC22K0487, HST grant HST-GO-15647.020-A, and the U.S. Department of Energy, Office of Science, Office of High Energy Physics Cosmic Frontier Research program under Award Number DE-SC0010008. The authors would also like to acknowledge guidance from Alice Shapley and Bahram Mobasher for guidance on the MOSDEF public data release.

REFERENCES

Anglés-Alcázar, D., Faucher-Giguère, C.-A., Quataert, E., et al. 2017, MNRAS Letters, 472, L109
Azadi, M., Coil, A., Aird, J., et al. 2018, ApJ, 866, 63
Azadi, M., Coil, A. L., Aird, J., et al. 2017, ApJ, 835, 27
Benítez-Llambay, A., Navarro, J. F., Abadi, M. G., et al. 2015, MNRAS, 450, 4207
Boogaard, L. A., Brinchmann, J., Bouché, N., et al. 2018, A&A, 619, A27
Brammer, G. B., van Dokkum, P. G., Franx, M., et al. 2012, ApJS, 200, 13
Brennan, R., Pandya, V., Somerville, R. S., et al. 2017, MNRAS, 465, 619
Broussard, A., Gawiser, E., Iyer, K., et al. 2019, ApJ, 873, 74
Byler, N., Dalcanton, J. J., Conroy, C., et al. 2017, ApJ, 840, 44
Calzetti, D., Armus, L., Bohlin, R. C., et al. 2000, ApJ, 533, 682
Chabrier, G. 2003, ApJL, 586, L133
Civano, F., Elvis, M., Brusa, M., et al. 2012, ApJS, 201, 30
Coil, A. L., Aird, J., Reddy, N., et al. 2015, ApJ, 801, 35
Conroy, C., & Gunn, J. E. 2010, ApJ, 712, 833
Conroy, C., Gunn, J. E., & White, M. 2009, ApJ, 699, 486
Daddi, E., Cimatti, A., Renzini, A., et al. 2004, ApJ, 617, 746
Daddi, E., Dickinson, M., Morrison, G., et al. 2007, ApJ, 670, 156
Davé, R., Rafieferantsoa, M. H., Thompson, R. J., et al. 2017, MNRAS, 467, 115
