

Balancing the efficiency and stochasticity of star formation with dust extinction in $z \gtrsim 10$ galaxies observed by JWST

Jordan Mirocha^{1*} and Steven R. Furlanetto²

¹*McGill University Department of Physics & McGill Space Institute, 3600 Rue University, Montréal, QC, H3A 2T8*

²*Department of Physics and Astronomy, University of California, Los Angeles, CA 90095, USA*

19 January 2024

ABSTRACT

Early observations with JWST indicate an over-abundance of bright galaxies at redshifts $z \gtrsim 10$ relative to *Hubble*-calibrated model predictions. More puzzling still is the apparent lack of evolution in the abundance of such objects between $z \sim 9$ and the highest redshifts yet probed, $z \sim 13$ – 17 . In this study, we first show that, despite a poor match with JWST LFs, semi-empirical models calibrated to UVLFs and colours at $4 \lesssim z \lesssim 8$ are largely consistent with constraints on the properties of individual JWST galaxies, including their stellar masses, ages, and rest-ultraviolet spectral slopes. We then show that order-of-magnitude scatter in the star formation rate of galaxies (at fixed halo mass) can indeed boost the abundance of bright galaxies, provided that star formation is more efficient than expected in low-mass halos. However, this solution to the abundance problem introduces tension elsewhere: because it relies on the up-scattering of low-mass halos into bright magnitude bins, one expects typical ages, masses, and spectral slopes to be much lower than constraints from galaxies observed thus far. This tension can be alleviated by non-negligible reddening, suggesting that – if the first batch of photometrically-selected candidates are confirmed – star formation *and* dust production could be more efficient than expected in galaxies at $z \gtrsim 10$.

Key words: galaxies: high-redshift – galaxies: luminosity function, mass function – diffuse radiation

1 INTRODUCTION

JWST has long promised a revolution in our understanding of early galaxy formation, and with the public release of the first deep surveys in July 2022, that revolution is finally beginning.

Several early analyses have extended studies of high-redshift galaxies to $z \gtrsim 10$ – and perhaps as far as $z \sim 17$ – using early data primarily taken from the GLASS-JWST and CEERS surveys. For the most part, these early searches have focused on photometric Lyman-break selection, often checking candidates with SED-fitting codes to attempt to validate the redshift selection. These studies have extended estimates of the UV luminosity function to much earlier times than heretofore possible (Donnan et al., 2023; Harikane et al., 2022).

Because of the limited depth of the early data, the most compelling candidates in these searches are relatively luminous (and hence presumably massive), and the biggest

surprise has been the apparent abundance of such massive galaxies at $z \gtrsim 10$, if the candidates are indeed primeval sources. Although the number counts are still small – and so heavily affected by Poisson noise and cosmic variance – they are consistent with the number density of bright galaxies remaining roughly constant at $z \gtrsim 8$ (Naidu et al., 2022b; Donnan et al., 2023; Castellano et al., 2022; Atek et al., 2022; Harikane et al., 2022; Labbe et al., 2022). This is very surprising in comparison to theoretical models, as the abundance of massive dark matter haloes should evolve very rapidly during this epoch (Mason et al., 2022; Boylan-Kolchin, 2022; Lovell et al., 2023).

These sources must be interpreted with substantial caution, however. Many of the faint sources are sensitive to the selection criteria (e.g., see the discussion in Harikane et al. 2022), while the recent release of in-flight photometric calibration corrections will drive $\sim$ tens of per-cent level changes in previously published photometry. Even bright sources can be confused by “low”-redshift interlopers. For example, Zavala et al. (2022) and Naidu et al. (2022a) show that dropout galaxies selected to lay at $z \sim 17$ may be con-

* jordan.mirocha@mcgill.ca

fused with dusty, star-forming galaxies at $z \sim 5$. Indeed, Glazebrook et al. (2022) and Rodighiero et al. (2023) find an array of galaxies with surprising properties, e.g., extreme obscuration ($A_V \sim 5$ at $z \sim 5$) and very strong Balmer breaks and emission lines at $2 \lesssim z \lesssim 6$. Given the extreme rarity of massive, high- z galaxies, reliable candidate selection requires comprehensive libraries of template spectra against which to compare (Furlanetto & Mirocha, 2022a) – a process made all the more difficult as JWST reveals the diversity of galaxy properties in the early Universe (e.g., Barrufet et al. 2022; Whitler et al. 2022).

Despite these caveats, it is intriguing to consider the implications of an “overabundance” of massive galaxies at $z \gtrsim 10$ – a phenomenon first suggested by the spectroscopically confirmed GN-z11 at $z \sim 11$ discovered in *Hubble Space Telescope* (HST) data (Oesch et al., 2016). Theoretical models robustly predict a sharp decline in the abundance of such objects, simply because the halo mass function is evolving rapidly at these times. Indeed, Boylan-Kolchin (2022) and Lovell et al. (2023) point out that the most extreme detections (from Labbe et al. 2022) appear to have much larger stellar masses than allowed by the standard cosmological paradigm, even if they have perfectly efficient star formation! Mason et al. (2022) show that the problem is less extreme for most sources but demonstrate that there is still a clear overabundance relative to the simplest models.

Previous work has shown that, through the HST era, the $z \gtrsim 6$ luminosity function could be understood using the same basic mechanisms that explain the abundance of star-forming galaxies at lower redshifts, with accretion onto haloes and stellar feedback controlling the star formation rates (e.g., Trenti et al. 2010; Mason et al. 2015; Furlanetto et al. 2017; Tacchella et al. 2018). Such “minimalist” models offered simple but powerful ways to extrapolate our understanding of galaxy formation to early times and roughly matched the predictions from more complex models (e.g., Behroozi & Silk 2015; Yung et al. 2019). But now it appears these kinds of models may significantly underpredict the abundance of massive sources, indicating that new physics is required to understand the first stages of galaxy formation – an exciting conclusion indeed!

So far, three mechanisms have been proposed to explain the overabundance. Mason et al. (2022) argue that it may be a result of scatter in the mapping of halo mass to UV luminosity – or in other words through strong fluctuations in the star formation rate of galaxies at a fixed halo mass. Because there are so many small haloes, galaxies with upward (temporary) fluctuations in the star formation rate will appear as luminous sources. The steepness of the underlying mass function means there are many more intrinsically small sources that will “up-scatter” to high luminosities than massive galaxies that will “down-scatter” to fainter luminosities, flattening the luminosity function and amplifying the bright end. One would also find an overabundance of bright objects if galaxies in low-mass halos produce stars or UV photons more efficiently than expected (Inayoshi et al., 2022).

Another potential explanation is dust. Measurements of dust obscuration in $z \sim 7$ galaxies by the REBELS survey show that a substantial fraction of star formation is obscured even at that time (Inami et al., 2022; Algera et al., 2023). Ferrara et al. (2022) argued that the overabundance of bright sources at $z \sim 10$ could be explained by evolu-

tion in the dust content: if very little obscuration occurred at early times (because the dust was not yet in place), the evolving dust content could roughly cancel the evolution in the halo mass function, causing the abundance of luminous galaxies to evolve only slowly.

In this work, we take a deeper look at the implications of an overabundance of luminous, high- z galaxies for our understanding of galaxy formation and evolution. In particular, we focus not just on the number counts themselves but on the inferred properties of these sources, which offer tantalizing clues about the mechanisms that may power these sources. In particular, we consider stellar mass estimates, inferred ages, and colours. We show that this additional information substantially complicates models that attempt to explain the overabundance of bright sources, requiring a linked set of non-trivial changes to the physics driving galaxy formation at early times.

In section 2, we describe the flexible framework through which we model the galaxy population and introduce the data sets to which we focus our comparison. We then show that models with typical assumptions cannot reasonably explain the new observations (§3) and that small number statistics arguments do not easily solve the abundance problem (§4). We then study several extensions to the model in §5, and we offer predictions for how each scenario’s assumptions manifest in high- z UVLFs and galaxy properties. We show that the wide range of measurements available with JWST will allow us to test these explanations in the near future. Finally, we conclude in section 6.

We adopt AB magnitudes throughout (Oke & Gunn, 1983), and the following cosmology: $\Omega_m = 0.3156$, $\Omega_b = 0.0491$, $h = 0.6726$, and $\sigma_8 = 0.8159$, similar to the recent Planck Collaboration et al. (2020) constraints.

2 METHODS

2.1 Galaxy models

We employ a semi-empirical model calibrated to UVLFs at $4 \lesssim z \lesssim 8$ and colours described in Mirocha et al. (2017, 2020), and implemented in the ARES code¹.

The basic approach is similar to other models in the recent literature, in that the star formation rate (SFR) of galaxies is assumed to be proportional to the mass accretion rate (MAR) of dark matter halos, modulo an efficiency factor that varies with redshift and/or halo mass (e.g., Trenti et al. 2010; Mason et al. 2015; Tacchella et al. 2018). We first initialize a set of halo assembly histories following the Furlanetto et al. (2017) approach, which assumes halos evolve at fixed number density and has been shown to agree well with the mean growth histories of halos in N -body simulations (Trac et al., 2015; Mirocha et al., 2021). We assume that galaxies occupy halos in a 1:1 fashion and that the halo mass function is that given by Tinker et al. (2010), which we compute using the HMF² code (Murray et al., 2013). We assume that the baryonic mass accretion rate is the total MAR times the cosmic baryon fraction, i.e., $\dot{M}_b = f_b \dot{M}_h$,

¹ <https://github.com/mirochaj/ares>

² <https://hmf.readthedocs.io/en/latest/>

where $f_b \approx 0.16$ is the cosmic baryon fraction. For our fiducial models, we assume 0.3 dex log-normal scatter in the MAR at fixed M_h , which matches measurements in cosmological simulations (Ren et al., 2019; Mirocha et al., 2021), but do not allow any other scatter in the star formation process. We will explore scenarios with additional scatter in §5.

With halo growth histories in hand, the SFR is then taken to be $\dot{M}_* = f_* \dot{M}_b$, where the star formation efficiency f_* is assumed to be a double-power law in halo mass,

$$f_*(M_h) = \frac{f_{*,10} \mathcal{C}_{10}}{\left(\frac{M_h}{M_p}\right)^{-\alpha_{*,lo}} + \left(\frac{M_h}{M_p}\right)^{-\alpha_{*,hi}}} \quad (1)$$

Here, $f_{*,10}$ is the SFE at $10^{10} M_\odot$ (which may be redshift-dependent), M_p is the mass at which f_* peaks, and $\alpha_{*,hi}$ and $\alpha_{*,lo}$ describe the power-law index at masses above and below the peak, respectively. The constant $\mathcal{C}_{10} \equiv (10^{10}/M_p)^{-\alpha_{*,lo}} + (10^{10}/M_p)^{-\alpha_{*,hi}}$ re-normalizes the standard double-power law formula to $10^{10} M_\odot$ instead of the peak mass. We also allow star formation to be stochastic, with high- z galaxies going through periods of star formation (with probability f_{duty} at any given time step) and dormancy.

We model the build-up of dust in galaxies assuming that the metal production rate, $\dot{M}_Z$, is proportional to the SFR, and that the dust yield is a fraction f_{dtr} of the metal yield. For an effective dust scale length, R_d , the dust optical depth can be written

$$\tau_\nu = \kappa_\lambda N_d = \kappa_\lambda \frac{3f_{\text{dtr}} M_Z}{4\pi R_d^2} \quad (2)$$

where we take the dust opacity to be $\kappa \propto \lambda^{-1}$, similar to SMC-like attenuation curves in the rest-UV (Weingartner & Draine, 2001). We parameterize R_d as a double power-law in halo mass, whose parameters (which may vary with redshift) are calibrated empirically jointly with the parameters of f_* via fits to UVLFs at $4 \lesssim z \lesssim 8$ (from Bouwens et al., 2015) and colours (from Bouwens et al., 2014). The full rest-ultraviolet spectrum of each halo in the model is synthesized over its past star formation history, including nebular continuum emission (following standard procedures Ferland, 1980), as implemented within *ARES* (Mirocha et al., 2020; Sun et al., 2021). This intrinsic spectrum is then reddened by the optical depth given in Eq. (2). We quantify their colours in terms of the spectral index β , where the rest-UV specific flux density is approximated as $f_\lambda \propto \lambda^\beta$.

We will explore two previously-established versions of this model to begin, both of which match UVLFs measured by HST, followed in section 5 by several extensions that aim to understand the implications of early JWST constraints. We take as our first case a modified version of the Mirocha et al. (2020) model, which used UVLFs at $z \sim 4, 6$, and 8, and colours at $z \sim 4$ and 6 (from Bouwens et al., 2015, 2014) to calibrate the free parameters of f_* (assumed independent of redshift here) and R_d . Here, for simplicity we neglect dust, but still synthesize the full SED of all model galaxies to establish a baseline for their rest-UV colours. Second, we employ a more physically-motivated model, in which f_* is given by the scalings appropriate for energy-regulated stellar feedback, $f_* \propto M_h^{2/3}(1+z)$ (as in Furlanetto et al., 2017), and the dust scale length evolves as the halo virial radius, $R_d \propto M_h^{1/3}(1+z)^{-1}$. In this case, reconciling the model

with UVLFs and colours requires evolution in both the star formation duty cycle (f_{duty}) and dust-to-metal-ratio (f_{dtr}) (Mirocha, 2020), because the shallow mass dependence of R_{vir} results in a dramatic over-reddening of the bright galaxy population (see also, e.g., Somerville et al., 2012; Yung et al., 2019).

As we will see shortly, neither of these models describe all of the new JWST observations as-is – both under-predict the $z \gtrsim 10$ galaxy abundance – which is unsurprising given that our approach is similar to others in the literature already known to be in tension with the new observations (as summarized in, e.g., Fig. 5 of Naidu et al., 2022b). As a result, we explore several variants of the model in what follows, in an attempt to identify viable scenarios to account for the new observations. While our approach is admittedly phenomenological, it has the virtue of making testable predictions for other galaxy observables.

2.2 Comparison data

The early release of JWST data has been accompanied by a bevy of observational analyses. We focus our comparison in this work on a subset of these, largely because they present either UVLFs (Donnan et al., 2023; Harikane et al., 2022; Naidu et al., 2022b, e.g.) or a key set of galaxy observables inferred from SED fitting, especially stellar masses, UV colours, and ages of individual objects (as reported in Atek et al., 2022; Whitler et al., 2022; Naidu et al., 2022b; Donnan et al., 2023; Harikane et al., 2022; Morishita & Stiavelli, 2022; Cullen et al., 2022; Finkelstein et al., 2022b).

These studies use three independent survey volumes. The GLASS-JWST survey uses a single JWST pointing, which is 9.7 arcmin^2 (Treu et al., 2022). CEERS is a somewhat larger survey, which so far has four pointings covering $\sim 33 \text{ arcmin}^2$ (Finkelstein et al., 2022a). Finally, the SMACS J0723 field has a small volume (a single pointing of 9.7 arcmin^2) but uses a foreground galaxy cluster to magnify distant sources (Ebeling et al., 2010; Coe et al., 2019; Ferreira et al., 2022).

Several other groups have made similar measurements of galaxies in this era: for example, Castellano et al. (2022) found several candidate galaxies at $9 \lesssim z \lesssim 15$ in the GLASS field, and Furtak et al. (2022) focused on strongly-lensed galaxies at $z \sim 10$. Yan et al. (2022) used color selection to develop a large list of candidates but did not attempt to measure physical quantities from them, while Carnall et al. (2023), Adams et al. (2023), and Trussler et al. (2022) focused on the detailed properties of a small number of objects. However, because all these studies use a subset of the same fields, they cannot be treated as fully independent measurements of the galaxy abundance. We have therefore focused on the aforementioned subset for convenience, recognizing that the detailed list of candidate galaxies depends upon the analysis method (see the helpful comparison in Harikane et al. 2022). However, the apparent overabundance of luminous galaxies at $z \gtrsim 8$ appears to be generic to all these analyses.

We also note that several studies have focused on more detailed properties of galaxies at $z \lesssim 8$ (e.g., Dressler et al. 2022). While eventually it will be crucial to match all of the properties of $z \gtrsim 10$ galaxies onto their descendants (and to use faint analogs at $z \lesssim 7$ to inform expectations for their

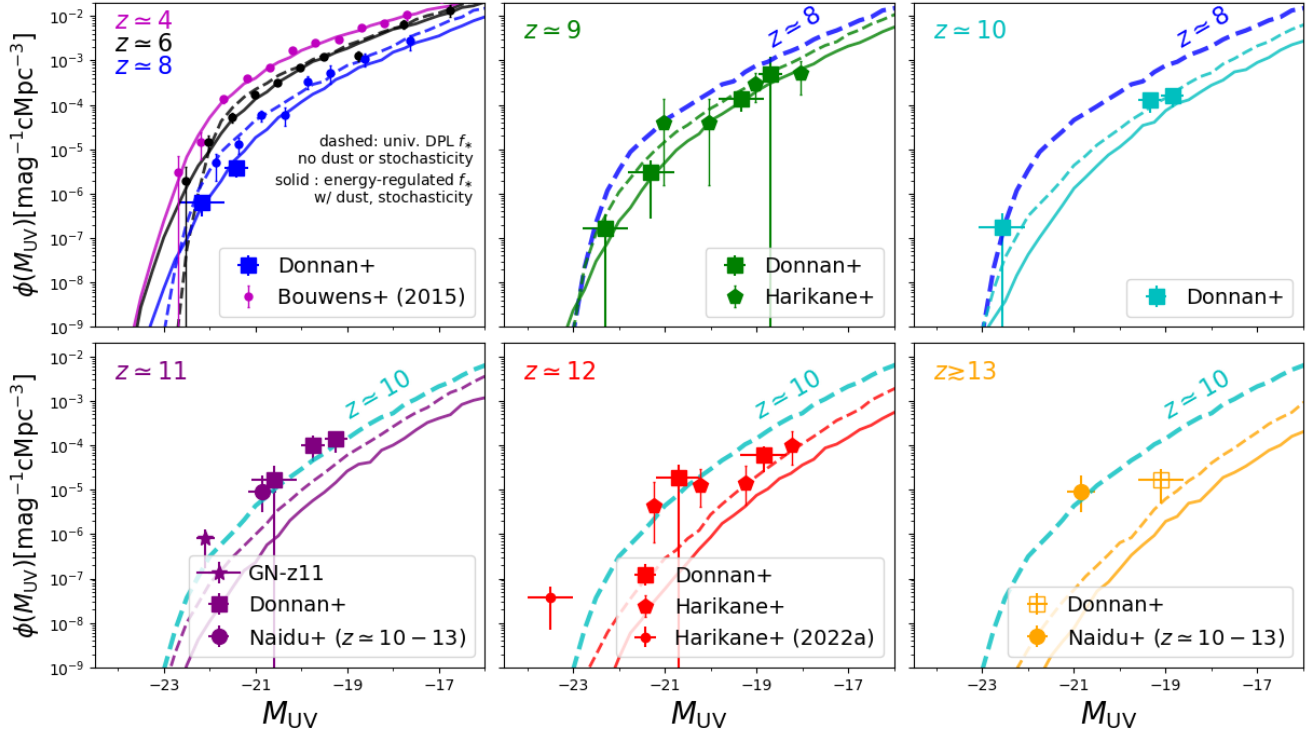


Figure 1. Rest-ultraviolet luminosity functions at high redshifts. Upper left panel indicates UVLFs assembled with Hubble at $z \sim 4, 6$, and 8 (Bouwens et al., 2015), with two additional points at $z \sim 8$ from JWST (Donnan et al., 2023). Remaining panels step through redshifts, as indicated in the upper-left corner of each panel. Dashed lines show the predictions of a model with a universal double power-law model for f_* and no dust attenuation, while the solid lines indicate the energy-regulated stellar feedback model of Mirocha (2020), which requires evolution in both f_{duty} and f_{dtnr} to satisfy UVLFs and colours simultaneously. Models like these systematically under-predict the abundance of bright $z \gtrsim 10$ sources revealed with JWST. Note that data points are not unique, i.e., the samples of galaxies assembled in recent studies are partially-overlapping (Naidu et al., 2022b; Donnan et al., 2023; Harikane et al., 2022). To that end, we mark measurements that do not appear in the most conservative of the colour cuts (from Harikane et al. 2022) with open symbols.

higher- z cousins through such studies as Nanayakkara et al. (2022), here we focus on the properties of these early galaxies inferred directly from their detections.

In this work, we will also take the properties of these galaxies inferred through SED-fitting – especially the stellar mass and age – at face value. This may be problematic, as both quantities depend upon the assumed parameterization of the star formation (see especially Whittler et al. 2022) and upon assumptions about the star formation process (e.g., Steinhardt et al. 2022), so difficulties in matching the observations may ultimately point toward refinements needed in such inferences. In particular, halos hosting galaxies grow exponentially with redshift during this era in essentially every theoretical model because halo growth rates are roughly proportional to halo mass (e.g., McBride et al., 2009; Goerdt et al., 2015), so we expect the star formation history to at least roughly mirror this behavior (see, e.g., Dekel & Mandelker, 2014; Furlanetto, 2021). A fair comparison to galaxy models requires star formation histories that permit this form in the inference procedure, though that is not generally included. We note that these different star formation histories make the age comparison particularly fraught.

Finally, we emphasize that even with the subset of measurements we have chosen to focus on, they all still ultimately draw from the same survey fields. Candidate galaxies

are either detected by more than one study (in which case there may be conflicting measurements of the same source) or are detected by one study but not another. To highlight the latter, we use Harikane et al. (2022), which appears to have the most conservative selection criteria, to highlight sources that are more marginal.

3 STATUS OF CURRENT MODELS

The most obvious challenge posed by early JWST observations, as pointed out previously (e.g., Naidu et al., 2022b; Mason et al., 2022; Ferrara et al., 2022), is to explain the apparent over-abundance of bright $z \gtrsim 10$ galaxies relative to model predictions anchored to $z \lesssim 8$ datasets. In this section, we start by showing the predictions of our default semi-empirical models from $z \sim 4$ to $z \sim 13$ for the UVLF (§3.1 and Fig. 1), and then focus on the same models’ predictions for the properties of individual galaxies (§3.2 and Fig. 2).

3.1 The UV luminosity function at $z \gtrsim 10$

In Fig. 1, we show UVLF predictions from our models at $z \sim 4 - 13$. UVLFs at $4 \lesssim z \lesssim 8$ agree with the plotted

data (from [Bouwens et al., 2015](#)) by construction (top-left corner), as do the M_{UV} - β relations during that period (see below), but the higher-redshift data were not used for the calibration. The five panels at $z \gtrsim 9$ thus demonstrate how HST-era models fare against the new observations.

Again, we emphasize that while data points from JWST in each of these figures are drawn from different studies, the same galaxies will appear in multiple datasets, potentially with different inferred properties. We have made no attempt to eliminate repeated sources in this work, or to give preference to one study over any other. However, given that [Harikane et al. \(2022\)](#) impose the most conservative colour cuts, we use open plot symbols to mark galaxies that were identified in the indicated paper but *not* by [Harikane et al. \(2022\)](#). We will discuss uncertainties arising from these small survey areas in §4.

The top row of Fig. 1 shows that both of our fiducial models fare reasonably well at $z \lesssim 10$, despite some mild tension at $z \sim 10$ for the energy-regulated model (top right, solid line). However, neither capture the abundance of $z \gtrsim 11$ galaxies inferred in early JWST analyses. The JWST UVLFs are consistent with no evolution between $z \sim 10$ and $z \gtrsim 13$, at least in the brightest M_{UV} bins.

In contrast, both of our models predict an order-of-magnitude decrease in the abundance of bright galaxies from $z \sim 10$ –13. This is most clear upon comparing to the dust-free universal SFE model prediction for $z = 10$, which is repeated in each panel of the bottom row to guide the eye (and which fits the $z \sim 10$ data well). This is a simple consequence of the rapid evolution in the halo mass function over this period: galaxies this luminous live inside massive haloes, which are undergoing rapid assembly throughout the Cosmic Dawn. This is the same tension highlighted by [Naidu et al. \(2022b\)](#) and addressed previously by [Mason et al. \(2022\)](#) and [Ferrara et al. \(2022\)](#).

3.2 Galaxy scaling relations

We now step beyond the abundance of these sources and ask another set of questions about how well the models describe the *properties* of these galaxies – in particular, the relationships between UV magnitude M_{UV} , stellar mass, stellar age, and ultraviolet colour³, all shown in Figure 2. We note that the M_{UV} - β relation at $z \sim 4$ and 6 was used to calibrate the model parameters so agrees with the plotted data from [Bouwens et al. \(2014\)](#) by construction. However, the stellar masses and ages are predictions at all redshifts, as these quantities have not been used at any stage of the model calibration. The observed M_{UV} - β relation at $z \gtrsim 8$ is also a prediction of the model. We do not show the M_* - β relation explicitly, but the [Mirocha et al. \(2020\)](#) model agrees well with the measurements of [Finkelstein et al. \(2012\)](#) at $4 \lesssim z \lesssim 8$.

In Fig. 2, we explore scaling relations between M_{UV} and M_* (top row), UV colour β (middle row), and age (bottom row), where in the models we define the age as the timescale

over which galaxies formed half their stellar mass. The half-mass time is slightly longer than mass-weighted ages (at least for exponential SFHs like ours), but of course shorter than the time since the first star formed, providing a middle-ground when comparing to the many observations in this study, in which each of these three definitions can be found. From left-to-right, we show our model predictions relative to observational constraints from $z \sim 4$ to $\gtrsim 13$. For clarity, we only include the level of scatter (as shaded regions) for the energy-regulated stellar feedback model, showing the median relations for the dust-free universal f_* model as dashed curves only.

We compare to constraints from a variety of recent studies ([Whitler et al., 2022](#); [Morishita & Stiavelli, 2022](#); [Atek et al., 2022](#); [Naidu et al., 2022b](#); [Finkelstein et al., 2022b](#); [Leethochawalit et al., 2022](#); [Cullen et al., 2022](#)), as indicated in the legends. Several SED-fitting codes are represented among these studies, which may infer different galaxy properties given the same input photometry. The inferred masses and ages, for example, can depend strongly on the assumed parameterization of the star formation history, whereas the colours are a more empirical diagnostic of galaxy measurements. [Whitler et al. \(2022\)](#) provide constraints on stellar population properties for a variety of assumptions – here, we adopt their results assuming a constant star formation rate with the BEAGLE code ([Chevallard & Charlot, 2016](#)). Non-parametric histories, e.g., from PROSPECTOR ([Johnson et al., 2021](#)), generally result in smaller ages (see also [Taccella et al., 2022](#)), which may help alleviate some of the tension in the M_{UV} -age plane. We do not attempt to add these systematic uncertainties to Figure 2, but we recommend the reader bear them in mind while comparing.

Both of our models predict similar M_{UV} - M_* and M_{UV} -age relationships (top and bottom rows, respectively), which are in good agreement with the latest JWST constraints, with the exception of a few very massive and old candidates at $z \sim 10$ (second column from left).

The most noticeable and consistent tension in Fig. 2 is that the colours of the dust-free model are far too blue, with $\beta \simeq -3$, in comparison to the measurements. This aspect of the model comes as no surprise, given the exponentially rising (and continuous) star formation histories of galaxies in our theoretical models: without dust, the UV luminosity is dominated by very young, massive stars, which have an intrinsic colour $\beta \sim -3$. While a few observed galaxies have extreme colours consistent with $\beta \sim -3$, most are closer to $\beta \sim -2$ or 2.5 (see also, e.g., [Topping et al., 2022](#)).

The colours of galaxies in our energy-regulated model are in much better agreement with these observations, indicating that the dependencies of f_{duty} and f_{dtr} on halo mass and redshift inferred from observations at $z \lesssim 8$ work reasonably well at $z \gtrsim 10$. There are some clear outliers, particularly at $z \sim 8$ and $z \sim 10$, but again, the agreement is reasonably good overall.

In this model, dust reddens the starlight, even at high redshifts. This highlights how the observed colours present an immediate challenge to one potential solution to the over-abundance problem – in which a rapid increase in dust reddening from $z \sim 10$ to $z \sim 7$ roughly cancels the underlying halo mass function evolution ([Ferrara et al., 2022](#)). Such a solution would require a different way to redden the galax-

³ The colour is quantified by a power-law fit to the rest-UV spectrum, i.e., $f_\lambda \propto \lambda^\beta$, which we estimate from a fit to mock photometry in the JWST medium filters.

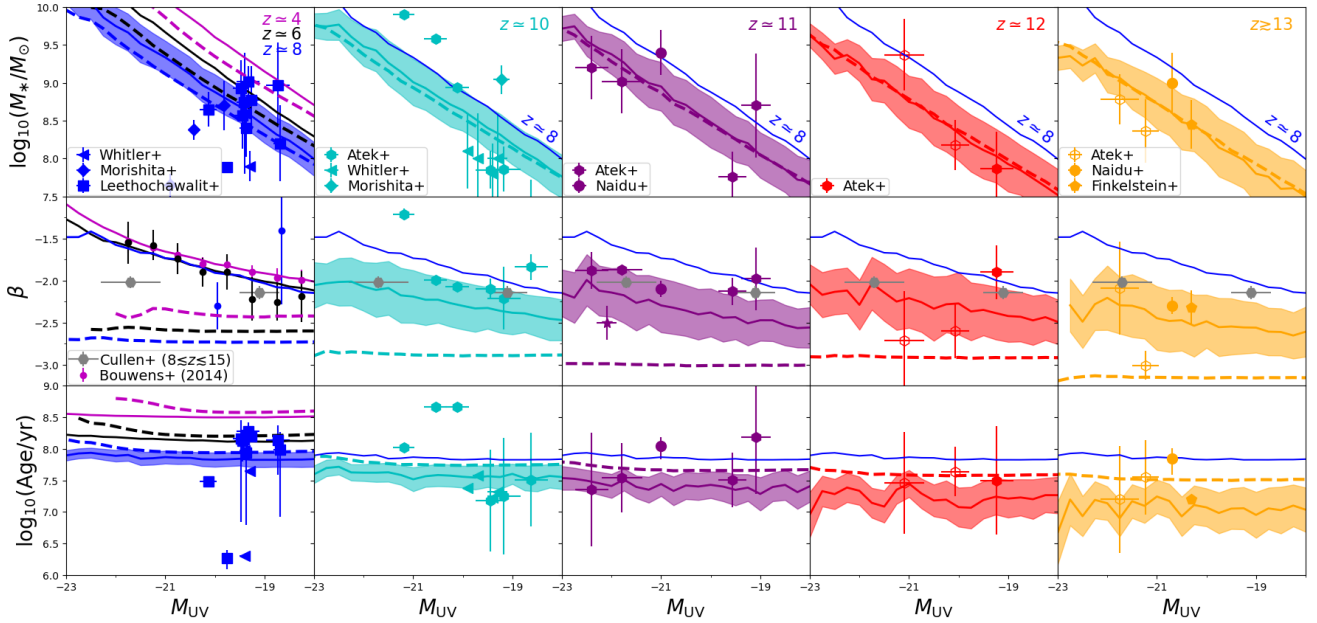


Figure 2. Stellar masses (top), rest-UV colours (middle), and age (bottom) as a function of M_{UV} from $z \sim 4$ to 13 (left to right). As in Fig. 1, dashed lines indicate the dust-free universal SFE model, while solid lines indicate the energy-regulated model with dust and burstiness. Shaded regions in the latter case indicate the level of intrinsic scatter predicted by the model, which arises due to a combination of 0.3 dex scatter in halo MAR and the stochasticity induced by $f_{duty} < 1$. A variety of datasets are included (see text). As in Fig. 1, open symbols are used to indicate galaxies that did not pass the Harikane et al. (2022) colour selection criteria.

ies, likely some form of quenching to ensure that they lack young, blue stars.

Given that the properties of individual galaxies are largely consistent with our model when dust is included, in the next section we explore the possibility that Poisson and cosmic variance could be responsible for the apparent over-abundance of $z \gtrsim 10$ galaxies.

4 POISSON SAMPLING AND COSMIC VARIANCE

The first wave of JWST results has been drawn from a few relatively small fields, so one might reasonably worry about the impact of small number statistics and cosmic variance. In Fig. 3, we explore these possibilities, with an emphasis on the probability of obtaining a $\sim$ constant number of galaxies over a wide redshift range.

In this section, we continue to use our fiducial models calibrated to the $4 \lesssim z \lesssim 8$ measurements, as in section 3. However, because the expected redshift evolution of these models is generic to many models based on the rapidly evolving halo mass function, we take a step back and consider how some basic properties of the halo population evolve over this period.

Starting in the top row, we show the predicted number of objects one expects to detect in a 10 arcmin² survey, of radial depth $\Delta z = 1$, comparable to the first GLASS parallel field (which uncovered one $z \sim 12$ –13 candidate; Castellano et al., 2022; Naidu et al., 2022b; Harikane et al., 2022) as a function of redshift. We generate 300 realizations of the halo population in each such volume, subject only to Poisson

noise from the limited volume. We consider separately the effect of cosmic variance momentarily.

We first focus solely on two halo properties: the halo mass M_h (left panel) and mass accretion rate (center panel) (which also has an intrinsic variation in the model). In both panels, we have chosen the lower mass and accretion rate thresholds to yield (on average) ~ 1 galaxy at $z \sim 12$, as found in the observations. The larger threshold in each panel was chosen to highlight the behavior at even higher masses. Shaded regions indicate the variation in source counts at each redshift above a specified mass or MAR cut (gray and blue shaded regions) with 68% (dark) and 95% (light) probability.

In the left panel, one can see clearly the rapid decline in the abundance of halos of a fixed mass, suggesting that a simple model in which M_{UV} traces M_h alone is not a viable solution to the JWST abundance problem: if a survey is sensitive enough to detect ~ 1 halo at $z \sim 12$, and if galaxy luminosity traces halo mass, it should find ~ 10 haloes at $z \sim 10$.

In the middle column of Fig. 3, we shift focus to the mass accretion rate: in our model, the star formation rate (and hence luminosity) is directly tied to accretion rather than halo mass, so this is a better proxy for luminosity. The abundance of haloes above a fixed MAR threshold evolves more slowly than the abundance above a fixed mass threshold. This is because the MAR is a strongly increasing function of redshift at fixed M_h (crudely, $\dot{M}_h/M_h \propto (1+z)^{5/2}$; Neistein & Dekel 2008; Dekel & Mandelker 2014), which helps to partially cancel the rapid evolution of the mass function itself. Here, we can see that a halo MAR cut of

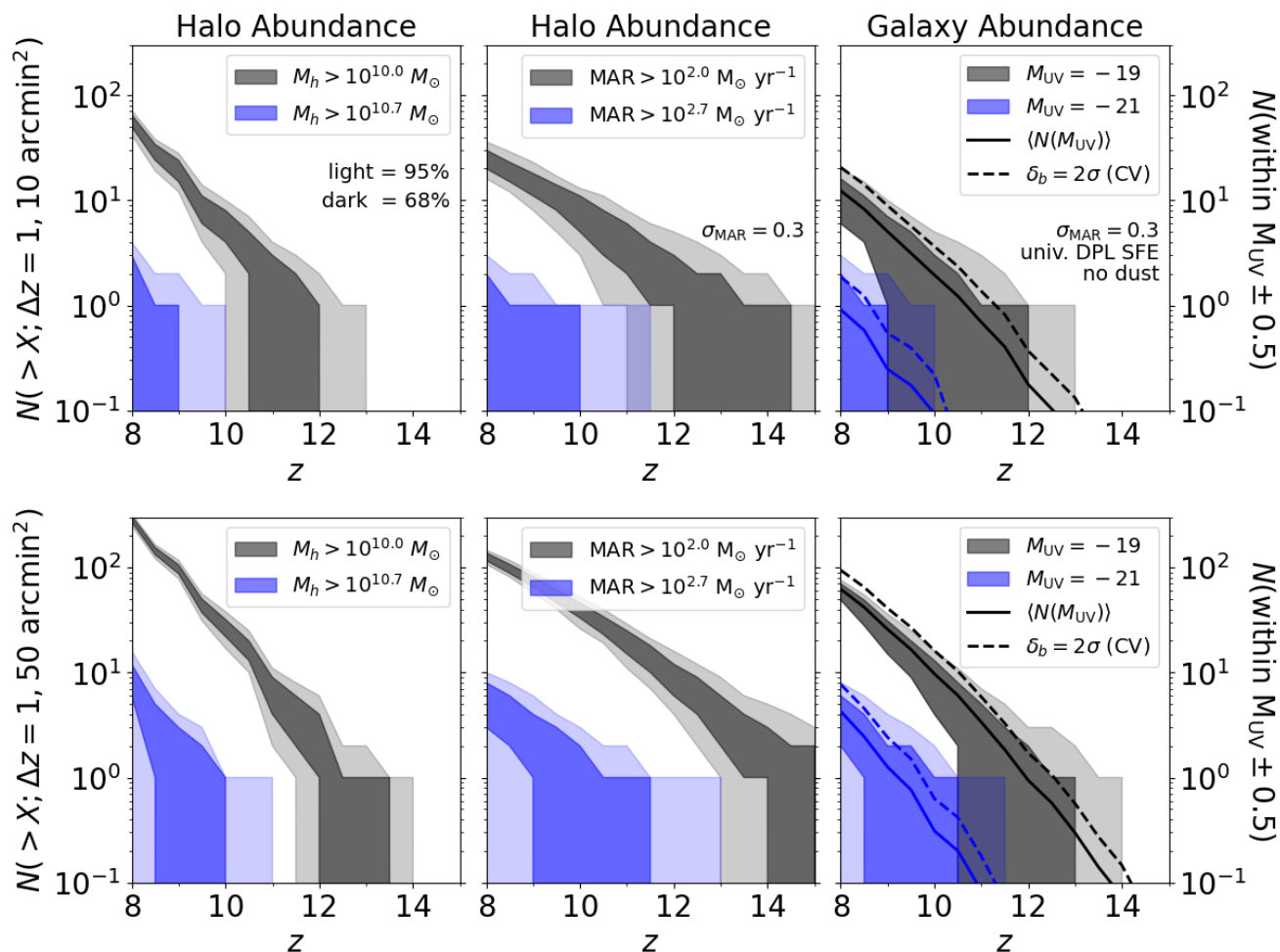


Figure 3. Abundance of objects above different halo mass thresholds (left), mass accretion rate thresholds (center), and within different M_{UV} bins (right). We assume $\Delta z = 1$ in a 10 arcmin² (top) and 50 arcmin² (bottom) field of view, comparable to the GLASS/SMACS field and total area of the first four CEERS fields, respectively, which have revealed $\sim 1 - 4$ galaxies per redshift bin at $z \gtrsim 10$. The rapid evolution of the HMF is clear in the left column, making it difficult to explain a $\sim$ constant abundance of galaxies at $z \gtrsim 10$, even in an idealized scenario in which M_h alone governs galaxy luminosity. Evolution in the counts at fixed MAR is more gradual, due to the strong redshift dependence in $MAR(M_h)$. However, typical galaxy models predict that f_* is a strong function of halo mass (right), making a constant abundance improbable. Dashed curves in the right column indicate the expected abundance of galaxies in a 2σ overdensity, while solid curves show the expected abundance at the cosmic mean density.

$\dot{M}_h \geq 10^{2.7} M_\odot \text{ yr}^{-1}$ yields $\sim 1 - 2$ halos 95% of the time between ~ 9 and ~ 11.5 .

However, one more step is needed to convert to UV luminosity. In most models, the efficiency of star formation is a strong function of halo mass (rather than MAR), which means that a fixed MAR threshold does *not* yield a fixed M_{UV} independent of redshift (because the $MAR-M_h$ relation evolves with redshift). We use our dust-free model to convert to UV luminosity in the final column of Figure 3, which shows the expected number of galaxies in two M_{UV} bins: the solid curve shows the mean expectation (matching the luminosity function shown in Fig. 1), while the shaded regions show the Poisson noise.

Given the small expected number of sources in each redshift bin, it is not surprising that Poisson sampling imposes a wide uncertainty on the source counts. In blue, we see that one expects 0 – 1 galaxies with $M_{UV} = -21$ over

$z \sim 9 - 10.5$. But even including the Poisson variations, the likelihood to find a source that bright at $z \gtrsim 11$ is extremely small. Thus Poisson noise alone is not enough to rescue our default models.

In the bottom row of Fig. 3, we perform the same exercise for a wider field of view, 50 arcmin², which is comparable to the total area of the first four CEERS fields (Finkelstein et al., 2022a). The number of targets increases, as expected, but the trends remain largely the same – the fluctuations in the counts of luminous sources are large, but not nearly large enough to explain the observations.

In the final column of Fig. 3, we also show examples of the effect of cosmic variance, accounting for the possibility that the survey fields happen to be overdense regions with enhanced halo populations. The mean expected number of galaxies in each magnitude bin is indicated by a solid line, while the dashed line indicates the boost in counts ex-

model	f_*	σ_{MAR}	dust yield	relevant figure
constant f_*	0.110	0.3	0.0	4
constant f_*	0.0066	1.0	0.0	4
constant f_*	0.150	0.3	0.1	5
constant f_*	0.072	1.0	0.125	5
DPL f_*	0.055	0.3	0.4	n/a

Table 1. Summary of model extensions shown graphically in Fig. 4 and 5. Note that these values are not determined via detailed fits. The key point is that (i) increased stochasticity (through σ_{MAR}) requires a reduction in the star formation efficiency to avoid over-producing galaxies, and (ii) the introduction of dust requires a slight increase in f_* to get roughly the same luminosity per SFR. We include for reference in the last row parameter values for the universal double power-law f_* model from Mirocha et al. (2020) that fits $4z \lesssim 8$ UVLFs and colours measured from *Hubble* data (Bouwens et al., 2014, 2015). In this case, the quoted f_* refers to the value at $M_h = 10^{10} M_\odot$.

pected in a region of the Universe that is a 2σ overdensity. These boosts are computed with *galcv* (Trapp & Furlanetto, 2020), an analytic model that assumes galaxy counts in different regions reflect differences in the underlying dark matter halo population alone, i.e., no additional variations in galaxy properties are introduced to reflect environmental effects on galaxy formation⁴. This overdensity provides a modest boost to the source counts, but it is far too small to explain the $\sim$ constant abundance of bright $M_{\text{UV}} \sim -21$ objects revealed with JWST, especially for wider area samples assembled in, e.g., Harikane et al. (2022).

We must also note another problem for cosmic variance and Poisson sampling explanations for the overabundance of $z \gtrsim 10$ sources: these excess sources appear across several redshift bins and multiple fields. While a statistical fluke could perhaps explain an overabundance in a single field and at a single redshift as an unlikely but plausible event, finding similar overabundances in many different volumes is extremely unlikely.

5 MODEL EXTENSIONS

From §3-4, it is clear that typical semi-empirical models calibrated to $z \lesssim 8$ datasets do not naturally explain the abundance of high- z galaxies discovered with JWST. In this section, we explore several extensions of the model in order to identify plausible explanations for the JWST data. We do not attempt to constrain quantitatively the physical elements of these solutions; instead, we illustrate the complexity of explaining the observations with a few specific examples that highlight the major challenges. We record the key parameter values of each model extension in Table 1 for reference.

⁴ Here we assume the galaxy counts follow a gamma distribution, which more accurately reflects cosmic variance in the rare-source, high-bias limit and has a longer tail for clustered sources (Steinhardt et al., 2021; Trapp et al., 2022).

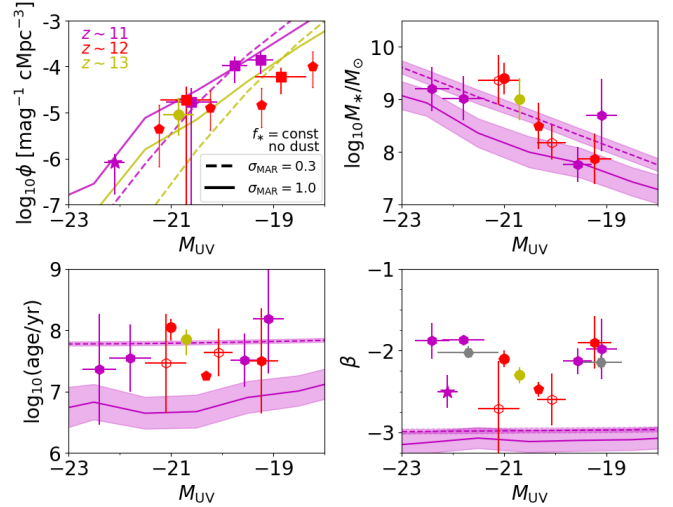


Figure 4. Effects of scatter in $\text{SFR}(M_h)$ on the abundance and properties of $z \gtrsim 10$ galaxies. Moving clockwise from the top left, the four panels show UVLFs, stellar masses, ages, and colours of galaxies at $z \sim 11, 12$, and 13 (magenta, red, and yellow lines/symbols, respectively). In the models shown here, we neglect dust, assume f_* is a constant, and vary the amount of scatter in $\text{SFR}(M_h)$ from 0.3 to 1 dex (dashed and solid lines, respectively). Data are the same as those shown in Fig. 1 and 2. Without dust, the rapidly rising star formation histories of the model galaxies predict very blue colours (lower right) and young ages (lower left). Note that model predictions at each redshift are not shown in every panel for clarity of presentation.

5.1 The star formation efficiency

We saw in Fig. 3 that the abundance of dark matter halos above a given MAR threshold can evolve quite slowly and plausibly produce $\sim 1 - 2$ objects over a broad range of redshifts $z \sim 9 - 13$, depending on the survey area. This same behaviour is not mirrored in the number counts of galaxies within a given UV magnitude bin in our models due to the strong halo mass dependence of the star formation efficiency (right column of Fig. 3).

We first therefore explore scenarios in which f_* is a constant, independent of both redshift and halo mass, which sets up a scenario in which galaxy star formation rates are directly proportional to halo mass accretion rates and may thus evolve relatively slowly with redshift to better match JWST constraints.

While we are motivated to fix f_* by this phenomenological result, we must emphasize that a constant star formation efficiency is a substantial departure from the results of galaxy models at lower redshifts. For example, the empirical models of Behroozi et al. (2019) demonstrate that the star formation efficiency of low-mass haloes at $z \lesssim 6$ increases with halo mass, while that of massive haloes decreases. The former effect is normally interpreted as “feedback regulation,” in which stellar radiation and/or supernovae prevents star formation and/or gas accretion in a way that depends upon the potential of the halo. If f_* is independent of halo properties, star formation occurs without knowledge of this larger environment and must depend only on local physics. One example of a process that can help “decouple” star formation from the halo environment is burstiness, which can

break the quasi-equilibrium relation between feedback and accretion (Orr et al., 2019; Furlanetto & Mirocha, 2022b). We note that Inayoshi et al. (2022) also considered a constant f_* scenario to explain the JWST observations, with an emphasis on the interplay between f_* and the efficiency of UV photon production. We will compare to their results momentarily.

In comparison to our fiducial models, a constant f_* boosts the efficiency of star formation in low-mass halos well above their nominal values and so inflates the pool of objects that could up-scatter into bright M_{UV} bins. (Here, we assume that scattering is a result of variations in the MAR, which follows a log-normal distribution with a scatter of $\sigma = 0.3$, as in the fiducial models.)

The result of this exercise is shown in Fig. 4. The four panels show our model extension’s predictions for UVLFs and relationships between M_{UV} and M_* , stellar age, and UV colour, going clockwise from top-left to bottom-left. For simplicity, we first focus on scenarios without dust. The dashed curves show our revised model with a constant $f_* = 0.11$, similar to the values found to roughly match the $z \sim 10$ JWST UVLFs in Inayoshi et al. (2022).

We find that in this model, tuned to roughly match the $z \sim 11$ UVLFs (magenta curves), the UVLF at $z \sim 13$ drops by more than an order of magnitude and cannot explain the Naidu et al. (2022b) or Donnan et al. (2023) sources (dashed yellow curves). Inayoshi et al. (2022) showed that $f_* \sim 0.3$ may be required to explain the $z \sim 13 - 17$ UVLFs, at least for normal stellar populations. Next, we will consider an alternative possibility: that f_* does not evolve over this interval but that the scatter in SFR at fixed halo mass is large.

5.2 Scatter in the halo–SFR relationship

We next explore the effect of increasing the scatter in the halo–galaxy mapping. To do so, we simply change the amount of log-normal scatter in the SFR of galaxies at fixed halo mass. Specifically, we increase the scatter from $\sigma = 0.3$, to $\sigma = 1$. This level of scatter has not been observed in models or simulations of dark matter halos at high redshift (Ren et al., 2019; Mirocha et al., 2021), but because $SFR \propto f_* \times MAR$ in our model, one can think of this scatter either as an increase in the MAR variations or as fluctuations intrinsic to the star formation process itself – perhaps in the star formation efficiency f_* or stochasticity in that process. To accommodate the increased stochasticity, we reduce the overall normalization of f_* from 0.11 to 0.006 to roughly preserve the LF at fainter magnitudes $M_{UV} \sim -19$. The large decrease in the *average* star formation efficiency is a manifestation of the importance of scatter, but we will see below it presents problems compared to the observations.

The solid curves in Figure 4 show the effects of this increased scatter. In this case, the abundance of bright objects is boosted significantly and the evolution of the UVLF with redshift slows accordingly, as suggested by Mason et al. (2022). The overall effect is to flatten the luminosity function at the bright end, because many more intrinsically small sources exist that can be up-scattered than massive sources that can be down-scattered. However, the resulting stellar masses and stellar ages are generally too small in comparison to the measured values. The latter is because up-scattered

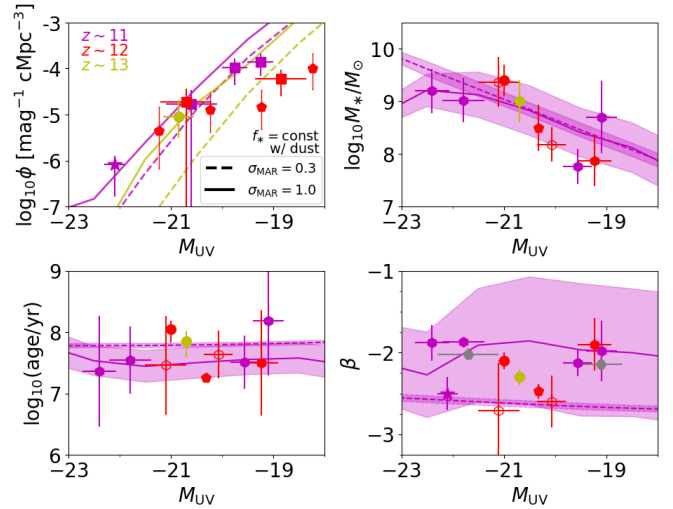


Figure 5. Effects of scatter in $SFR(M_h)$ and dust content on the abundance and properties of $z \gtrsim 10$ galaxies. Here, as in Fig. 4, we assume f_* is a constant and vary the amount of scatter in $SFR(M_h)$. However, we also allow for dust, which drives the model toward bluer colours, older ages, and higher stellar masses, in better agreement with early constraints from JWST.

objects are undergoing a temporary increase in their star formation rate so have very young stellar populations – in contrast to the measurements, where they are typically found to be ~ 100 Myr old. Similarly, the models underestimate the stellar masses because so much of the luminosity comes from the young stars – whereas older populations require more mass to make up the difference, and because the overall star formation efficiency is much smaller.

The most obvious problem, however, is with the UV colours: we have ignored dust, so the young stars driving luminous objects in the models have $\beta \simeq -3$, in severe tension with the JWST constraints (lower right panel).

5.3 The role of dust

Finally, in order to match the observed colors we re-introduce dust. As an example of how this affects the model results, we will explore the effects of changing the dust yield at $z \gtrsim 10$, and keep R_d fixed to a power-law with $R_d \propto M_h^{0.45}$ as in the fiducial model. We first tune f_* to achieve agreement in the M_{UV} – M_* relation (top right), then adjust the dust yield $f_{d, \text{tmr}}$ until the typical colour is $\beta \simeq -2$. For the $\sigma = 0.3$ case, we obtain $f_* = 0.15$ and $f_{d, \text{tmr}} = 0.1$. With $\sigma = 1$, we obtain $f_* = 0.07$ and $f_{d, \text{tmr}} = 0.125$. (Note that introducing dust increases the overall star formation efficiency by an order of magnitude in the high-scatter case, easing the tension with the stellar mass measurements.)

The results are shown in Figure 5, for both moderate and elevated scatter. Just as in the dust-free models, increasing the scatter slows the evolution in the UVLFs somewhat. However, the inclusion of dust partially counters the effects of scatter, making it more difficult to simultaneously match UVLFs and colours. The dust extinction increases with halo mass, acting more strongly on the most massive systems.

This steepens the mass function, making it more difficult to match both bright and fainter sources at high redshifts.

We note that the combination of increased (and constant) f_* , increased scatter, and dust presented here works reasonably well to match the observations but is by no means unique. For example, though we found $f_{\text{dtr}} \simeq 0.1$ – about four times smaller than the default $f_{\text{dtr}} = 0.4$ (motivated by, e.g., Dwek et al., 2007) – this parameter is degenerate with the normalization of the dust scale length (see eq. 2), a more difficult parameter to interpret or constrain. Additional degeneracies between the dust scale length and star formation parameters are also strong, and so other solutions may provide equally reasonable agreement. Their mutual dependencies should be explored more thoroughly in the future. Nevertheless, it is reassuring that the dust correction is relatively modest in comparison to models at later times, as dust typically forms over long timescales.

6 DISCUSSION & CONCLUSIONS

Motivated by the apparent over-abundance of bright sources at $z \gtrsim 10$ implied by the first wave of results from JWST, we use a semi-empirical modeling framework to identify galaxy evolution scenarios that can reproduce this excess. Importantly, we consider both the UVLFs and observed constraints on the properties of individual galaxies, including the stellar ages, stellar masses, and rest-ultraviolet colours.

We find that maintaining a constant number density of bright galaxies at $z \gtrsim 10$ requires several adjustments to “standard” *HST*-calibrated models. First, one must counteract the rapid decline of the abundance of massive haloes over this redshift range, e.g., by appealing to M_h -independent star formation efficiencies (see §4 and Fig. 3). This moderates, but does not eliminate, the expected decline in halo counts. Second, substantial scatter in galaxy SFRs at fixed M_h (over and above that found by variations in the mass accretion rate in cosmological simulations) improves the match to the shape of UVLFs and further slows the UVLF evolution. However, leveraging the abundance of low-mass haloes to boost counts in bright M_{UV} bins results in very young ages, low stellar masses, and blue colours that are inconsistent with JWST constraints on most individual galaxies (Cullen et al., 2022; Topping et al., 2022; Nanayakkara et al., 2022). As a result, we must also make a *third* change: non-trivial scatter must be accompanied by non-negligible dust, which helps not only to alleviate the tension with JWST colours, but also with stellar masses and ages as well, because dust extinction requires an increase in f_* to roughly preserve galaxy counts.

We must be cautious in interpreting these observations, of course: it has already been argued that some high- z candidates are interlopers from lower redshifts (Zavala et al., 2022; Naidu et al., 2022a), and there are good reasons to expect that contamination becomes a serious problem at $z \gtrsim 12$ (Furlanetto & Mirocha, 2022a). Moreover, we have taken inferences about stellar ages and masses from SED-fitting codes, which require an accurate set of template spectra (potentially a problem at high redshifts; Steinhart et al. 2022) and accurate parameterization of the star formation histories (e.g., imposing a constant star formation history will overestimate M_* compared to bursty models; Whitler

et al. 2022). Spectra of these candidates will be crucial in rejecting interlopers and more robustly measuring the stellar populations. Only then will we pin down the UVLFs at $z \gtrsim 10$; for now we can only make tentative conclusions.

But, provided that the candidates are confirmed, we emphasize that, while any one of our three solutions can be invoked to explain the apparent (nearly redshift-independent) overabundance of bright galaxies at $z \gtrsim 10$ (Mason et al., 2022; Ferrara et al., 2022; Inayoshi et al., 2022), it is much more difficult to create a scenario consistent with both the UVLFs and the properties of individual galaxies, even at this early stage of JWST observations. This is important because all three of our required changes are, in a sense, surprising:

(i) First, the steep dependence of f_* on M_h , which is essential to fitting observed UVLFs at $z \lesssim 8$, is generally attributed to stellar feedback. One interpretation of our results is that the JWST UVLFs support a picture in which stellar feedback is ineffective at $z \gtrsim 10$, so that f_* is independent of halo mass. This is not wholly unexpected; halo dynamical times are comparable to the lifetimes of massive stars at high redshifts (e.g., Faucher-Giguère, 2018; Orr et al., 2019; Furlanetto & Mirocha, 2022a), which can break the equilibrium established at later times between supernova-driven outflows and cosmological inflows. The qualitative expectation is for star formation to occur in bursts as a result of this disconnect. This is one example of a mechanism that breaks the connection between the star formation rate and the halo environment and so can cause a constant f_* – a fundamental change in the nature of galactic-scale star formation. We also require significantly more efficient star formation in these haloes than feedback arguments suggest; interestingly, the f_* values used here are comparable to those required to explain the EDGES detection of the global 21-cm signal (Bowman et al., 2018; Mirocha & Furlanetto, 2019).

(ii) We also require significantly more scatter in the $\text{SFR}(M_h)$ relation than previously expected (of about an order-of-magnitude, compared to 0.3 dex variation in the mass accretion rate of haloes measured by simulations). The excess scatter likely results from the star formation process itself, and it could either be attributed to short-term temporal variations in the SFR or stochasticity in the star formation process. This too is potentially consistent with a picture in which star formation occurs in bursts in this population, although of course other scenarios are possible.

(iii) Finally, an (almost) inevitable result of an increased scatter is to make luminous galaxies more blue (because their high luminosities are caused by very recent star formation episodes). In order to make the colours consistent with JWST measurements, we require non-negligible dust. This may itself be surprising at $z \gtrsim 10$, given the short timescales involved: it likely requires dust production in supernova explosions, because the systems are not old enough for dust to be produced in AGB stars. However, these systems are likely so compact that the total amount of dust does not need to be extreme, if it also remains compact. Indeed, our example scenario only requires dust production to be $\sim 25\%$ as efficient as in “normal” galaxies.

Our solution is not unique: degeneracies exist between these three factors, which should be explored in more detail in the future. For example, the colours also depend upon the ages of the stellar populations. Very young stellar populations with strong nebular continuum emission could

produce $\beta \sim -2.5$ colours even without dust. Similarly, stochastic star formation histories can result in colours redder than expected in galaxies that are between star formation episodes (see, e.g., [Kelson & Abramson, 2022](#)). The fact that our model – which includes nebular continuum emission and stochasticity – still requires dust to jointly fit number counts and colours could be a byproduct of our assumed parameterizations. For example, we implement stochasticity as a random ‘flickering’ in the SFR rather than as coherent bursts, which makes it very unlikely that all galaxies in a given M_{UV} bin are uniformly young and dust free. As a result, it may be misleading to compare the M_{UV} - β relation predictions for an entire galaxy population with the M_{UV} and β values for just a few individual galaxies.

Finally, we note that by focusing on the UVLF, we have not incorporated all the early JWST analyses. We have ignored candidates selected for their large stellar masses, such as those of [Labbe et al. \(2022\)](#), which pose substantial challenges to the standard galaxy evolution paradigm (as pointed out by, e.g., [Boylan-Kolchin, 2022](#); [Lovell et al., 2023](#)). However, strong nebular line emission could also bias inferred stellar masses significantly (see [Endsley et al., 2022](#)). Spectroscopic follow-up of such sources thus remains paramount.

At the least, this analysis has demonstrated the complexity of interpreting the early JWST measurements. We expect our understanding to evolve rapidly as larger and more refined samples become available in the coming months.

DATA AVAILABILITY

The data underlying this article is available upon request.

ACKNOWLEDGMENTS

We thank Adam Trapp for his help with `galcv`. SRF was supported by the National Science Foundation through award AST-1812458. In addition, this work was directly supported by the NASA Solar System Exploration Research Virtual Institute cooperative agreement number 80ARC017M0006. This work has made extensive use of NASA’s Astrophysics Data System (<http://ui.adsabs.harvard.edu/>) and the arXiv e-Print service (<http://arxiv.org>).

Software: `numpy` ([Van Der Walt et al., 2011](#)), `scipy` ([Virtanen et al., 2020](#)), `matplotlib` ([Hunter, 2007](#)), `hmf` ([Murray et al., 2013](#)).

References

- Adams N. J. et al., 2023, *MNRAS*, 518, 4755
 Algera H. S. B. et al., 2023, *MNRAS*, 518, 6142
 Atek H. et al., 2022, arXiv e-prints, arXiv:2207.12338
 Barrufet L. et al., 2022, arXiv e-prints, arXiv:2207.14733
 Behroozi P. et al., 2019, *MNRAS*, 488, 3143
 Behroozi P. S., Silk J., 2015, *ApJ*, 799, 32
 Bouwens R. J. et al., 2014, *ApJ*, 793, 115
 Bouwens R. J. et al., 2015, *ApJ*, 803, 34
 Bowman J. D. et al., 2018, *Nat*, 555, 67
 Boylan-Kolchin M., 2022, arXiv e-prints, arXiv:2208.01611
 Carnall A. C. et al., 2023, *MNRAS*, 518, L45
 Castellano M. et al., 2022, *ApJL*, 938, L15
 Chevallard J., Charlot S., 2016, *MNRAS*, 462, 1415
 Coe D. et al., 2019, *ApJ*, 884, 85
 Cullen F. et al., 2022, arXiv e-prints, arXiv:2208.04914
 Dekel A., Mandelker N., 2014, *MNRAS*, 444, 2071
 Donnan C. T. et al., 2023, *MNRAS*, 518, 6011
 Dressler A. et al., 2022, arXiv e-prints, arXiv:2208.04292
 Dwek E., Galliano F., Jones A. P., 2007, *ApJ*, 662, 927
 Ebeling H. et al., 2010, *MNRAS*, 407, 83
 Endsley R. et al., 2022, arXiv e-prints, arXiv:2208.14999
 Faucher-Giguère C.-A., 2018, *MNRAS*, 473, 3717
 Ferland G. J., 1980, *PASP*, 92, 596
 Ferrara A., Pallottini A., Dayal P., 2022, arXiv e-prints, arXiv:2208.00720
 Ferreira L. et al., 2022, *ApJL*, 938, L2
 Finkelstein S. L. et al., 2022a, arXiv e-prints, arXiv:2211.05792
 Finkelstein S. L. et al., 2022b, *ApJL*, 940, L55
 Finkelstein S. L. et al., 2012, *ApJ*, 756, 164
 Furlanetto S. R., 2021, *MNRAS*, 500, 3394
 Furlanetto S. R., Mirocha J., 2022a, *MNRAS*, submitted
 Furlanetto S. R., Mirocha J., 2022b, *MNRAS*, 511, 3895
 Furlanetto S. R. et al., 2017, *MNRAS*, 472, 1576
 Furtak L. J. et al., 2022, arXiv e-prints, arXiv:2208.05473
 Glazebrook K. et al., 2022, arXiv e-prints, arXiv:2208.03468
 Goerdt T. et al., 2015, *MNRAS*, 454, 637
 Harikane Y. et al., 2022, arXiv e-prints, arXiv:2208.01612
 Hunter J. D., 2007, Computing in Science & Engineering, 9, 90
 Inami H. et al., 2022, *MNRAS*, 515, 3126
 Inayoshi K. et al., 2022, *ApJL*, 938, L10
 Johnson B. D. et al., 2021, *ApJS*, 254, 22
 Kelson D. D., Abramson L. E., 2022, Research Notes of the American Astronomical Society, 6, 162
 Labbe I. et al., 2022, arXiv e-prints, arXiv:2207.12446
 Leethochawalit N. et al., 2022, arXiv e-prints, arXiv:2207.11135
 Lovell C. C. et al., 2023, *MNRAS*, 518, 2511
 Mason C. A., Trenti M., Treu T., 2015, *ApJ*, 813, 21
 Mason C. A., Trenti M., Treu T., 2022, arXiv e-prints, arXiv:2207.14808
 McBride J., Fakhouri O., Ma C.-P., 2009, *MNRAS*, 398, 1858
 Mirocha J., 2020, *MNRAS*, 499, 4534
 Mirocha J., Furlanetto S. R., 2019, *MNRAS*, 483, 1980
 Mirocha J., Furlanetto S. R., Sun G., 2017, *MNRAS*, 464, 1365
 Mirocha J., La Plante P., Liu A., 2021, *MNRAS*, 507, 3872
 Mirocha J., Mason C., Stark D. P., 2020, *MNRAS*, 498, 2645
 Morishita T., Stiavelli M., 2022, arXiv e-prints, arXiv:2207.11671
 Murray S. G., Power C., Robotham A. S. G., 2013, Astronomy and Computing, 3, 23
 Naidu R. P. et al., 2022a, arXiv e-prints, arXiv:2208.02794
 Naidu R. P. et al., 2022b, *ApJL*, 940, L14
 Nanayakkara T. et al., 2022, arXiv e-prints, arXiv:2207.13860
 Neistein E., Dekel A., 2008, *MNRAS*, 388, 1792

- Oesch P. A. et al., 2016, *ApJ*, 819, 129
- Oke J. B., Gunn J. E., 1983, *ApJ*, 266, 713
- Orr M. E., Hayward C. C., Hopkins P. F., 2019, *MNRAS*, 486, 4724
- Planck Collaboration et al., 2020, *A&A*, 641, A6
- Ren K., Trenti M., Mason C. A., 2019, *ApJ*, 878, 114
- Rodighiero G. et al., 2023, *MNRAS*, 518, L19
- Somerville R. S. et al., 2012, *MNRAS*, 423, 1992
- Steinhardt C. L., Jespersen C. K., Linzer N. B., 2021, *ApJ*, 923, 8
- Steinhardt C. L. et al., 2022, arXiv e-prints, arXiv:2208.07879
- Sun G. et al., 2021, *MNRAS*, 508, 1954
- Tacchella S. et al., 2018, *ApJ*, 868, 92
- Tacchella S. et al., 2022, arXiv e-prints, arXiv:2208.03281
- Tinker J. L. et al., 2010, *ApJ*, 724, 878
- Topping M. W. et al., 2022, arXiv e-prints, arXiv:2208.01610
- Trac H., Cen R., Mansfield P., 2015, *ApJ*, 813, 54
- Trapp A. C., Furlanetto S. R., 2020, *MNRAS*, 499, 2401
- Trapp A. C., Furlanetto S. R., Yang J., 2022, *MNRAS*, 510, 4844
- Trenti M. et al., 2010, *ApJL*, 714, L202
- Treu T. et al., 2022, *ApJ*, 935, 110
- Trussler J. A. A. et al., 2022, arXiv e-prints, arXiv:2207.14265
- Van Der Walt S., Colbert S. C., Varoquaux G., 2011, *Computing in Science & Engineering*, 13, 22
- Virtanen P. et al., 2020, *Nature Methods*, 17, 261
- Weingartner J. C., Draine B. T., 2001, *ApJ*, 548, 296
- Whitler L. et al., 2022, *MNRAS*
- Yan H. et al., 2022, arXiv e-prints, arXiv:2207.11558
- Yung L. Y. A. et al., 2019, *MNRAS*, 483, 2983
- Zavala J. A. et al., 2022, arXiv e-prints, arXiv:2208.01816