# The Foundation Supernova Survey: Measuring Cosmological Parameters with Supernovae from a Single Telescope

D. O. Jones[1], D. M. Scolnic[2,14], R. J. Foley[1], A. Rest[3,4], R. Kessler[2], P. M. Challis[5], K. C. Chambers[6], D. A. Coulter[1], K. G. Dettman[7], M. M. Foley[5], M. E. Huber[6], S. W. Jha[7], E. Johnson[4], C. D. Kilpatrick[1], R. P. Kirshner[5,8], J. Manuel[9], G. Narayan[3], Y.-C. Pan[10,11,15], A. G. Riess[3,4], A. S. B. Schultz[6], M. R. Siebert[1], E. Berger[5], R. Chornock[12], H. Flewelling[6], E. A. Magnier[6], S. J. Smartt[13], K. W. Smith[13], R. J. Wainscoat[6], C. Waters[6], and M. Willman[6]

[1] Department of Astronomy and Astrophysics, University of California, Santa Cruz, CA 92064, USA; david.jones@ucsc.edu
[2] University of Chicago, Kavli Institute for Cosmological Physics, Chicago, IL, USA
[3] Space Telescope Science Institute, Baltimore, MD 21218, USA
[4] Department of Physics and Astronomy, Johns Hopkins University, Baltimore, MD 21218, USA
[5] Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, MA 02138, USA
[6] Institute for Astronomy, University of Hawaii, 2680 Woodlawn Drive, Honolulu, HI 96822, USA
[7] Department of Physics and Astronomy, Rutgers, The State University of New Jersey, 136 Frelinghuysen Road, Piscataway, NJ 08854, USA
[8] Gordon and Betty Moore Foundation, 1661 Page Mill Road, Palo Alto, CA 94304, USA
[9] Department of Astrophysics and Planetary Science, Villanova University, Villanova, PA 19085, USA
[10] Division of Theoretical Astronomy, National Astronomical Observatory of Japan, 2-21-1 Osawa, Mitaka, Tokyo 181-8588, Japan
[11] Institute of Astronomy and Astrophysics, Academia Sinica, Taipei 10617, Taiwan
[12] Astrophysical Institute, Department of Physics and Astronomy, 251B Clippinger Lab, Ohio University, Athens, OH 45701, USA
[13] Astrophysics Research Centre, School of Mathematics and Physics, Queen's University Belfast, Belfast BT7 1NN, UK

## Abstract

Measurements of the dark energy equation-of-state parameter, $w$, have been limited by uncertainty in the selection effects and photometric calibration of $z < 0.1$ Type Ia supernovae (SNe Ia). The Foundation Supernova Survey is designed to lower these uncertainties by creating a new sample of $z < 0.1$ SNe Ia observed on the Pan-STARRS system. Here we combine the Foundation sample with SNe from the Pan-STARRS Medium Deep Survey and measure cosmological parameters with 1338 SNe from a single telescope and a single, well-calibrated photometric system. For the first time, both the low-$z$ and high-$z$ data are predominantly discovered by surveys that do not target preselected galaxies, reducing selection bias uncertainties. The $z > 0.1$ data include 875 SNe without spectroscopic classifications, and we show that we can robustly marginalize over CC SN contamination. We measure Foundation Hubble residuals to be fainter than the preexisting low-$z$ Hubble residuals by $0.046 \pm 0.027$ mag (stat + sys). By combining the SN Ia data with cosmic microwave background constraints, we find $w = -0.938 \pm 0.053$, consistent with ΛCDM. With 463 spectroscopically classified SNe Ia alone, we measure $w = -0.933 \pm 0.061$. Using the more homogeneous and better-characterized Foundation sample gives a 55% reduction in the systematic uncertainty attributed to SN Ia sample selection biases. Although use of just a single photometric system at low and high redshift increases the impact of photometric calibration uncertainties in this analysis, previous low-$z$ samples may have correlated calibration uncertainties that were neglected in past studies. The full Foundation sample will observe up to 800 SNe to anchor the LSST and *WFIRST* Hubble diagrams.

*Key words:* cosmology: observations – dark energy – supernovae: general

## 1. Introduction

Since the discovery of dark energy 20 years ago (Riess et al. 1998; Perlmutter et al. 1999), measurements of the dark energy equation-of-state parameter, $w$, have been steadily improving (Garnavich et al. 1998; Knop et al. 2003; Tonry et al. 2003; Riess et al. 2004, 2007; Astier et al. 2006; Wood-Vasey et al. 2007; Kowalski et al. 2008; Kessler et al. 2009; Sullivan et al. 2011; Betoule et al. 2014; Scolnic et al. 2018). In support of a better understanding of dark energy, recent cosmic microwave background (CMB) experiments have yielded improved measurements of the cosmic matter density at $z \approx 1090$ (Planck Collaboration et al. 2018), and baryon acoustic oscillations (BAO) have given excellent constraints on the acoustic scale from $z \approx 0.3$ to $z \approx 2$ (Anderson et al. 2014; Ross et al. 2015; Alam et al. 2017). As Type Ia supernova (SN Ia) sample sizes have steadily increased, their systematic uncertainties have steadily decreased. Their reduced systematic uncertainties are primarily due to improvements in photometric calibration and a better understanding of the ways in which SN Ia distance measurements are biased by selection effects. Systematic and statistical uncertainties on $w$ have been approximately equal in most recent measurements (e.g., Betoule et al. 2014; Jones et al. 2018b, hereafter J18; Scolnic et al. 2018, hereafter S18; Brout et al. 2019).

Counterintuitively, many of the dominant sources of systematic uncertainty in dark energy measurements stem from the nearest SNe Ia. While high-$z$ SN Ia samples from the Sloan Digital Sky Survey (SDSS; Kessler et al. 2009), the Supernova

---

[14] Hubble, KICP Fellow.
[15] EACOA Fellow.

Legacy Survey (SNLS; Conley et al. 2011; Sullivan et al. 2011), and Pan-STARRS (PS1; Rest et al. 2014; Scolnic et al. 2014b; S18; J18) are observed on photometric systems with mmag-level systematic uncertainties, the heterogeneous low-$z$ SN samples are observed on more than 13 different photometric systems, each with their own systematic uncertainties. These systematic uncertainties may be correlated in ways that are difficult to predict. There could be additional unknown systematic uncertainties associated with the fact that many of the data were taken at a time when cosmological analyses were not yet concerned with or limited by mmag-level systematic uncertainties. The sample selection criteria and follow-up criteria for low-$z$ SNe Ia are heterogeneous and sometimes not well documented, leading to systematic uncertainties in the observational biases and selection effects.

Unlike the high-$z$ SN Ia data, samples of SNe Ia at $z < 0.1$ were predominantly found by surveys that targeted preselected sets of galaxies (e.g., the Lick Observatory Supernova Search; Filippenko et al. 2001). There is some evidence that SNe Ia selected by targeted surveys have different biases than those from untargeted surveys (Jones et al. 2018a). The Foundation Supernova Survey (Foley et al. 2018) aims to create a single, low-$z$ sample that is more similar in calibration uncertainty and sample selection characteristics to the high-$z$ data. Foundation uses the PS1 telescope to follow SNe Ia discovered primarily by untargeted searches such as ASAS-SN (Holoien et al. 2017), ATLAS (Tonry et al. 2018), *Gaia* (Gaia Collaboration et al. 2016), and the Pan-STARRS Survey for Transients (PSST; Huber et al. 2015). Although some Foundation data were discovered by targeted searches, untargeted surveys would likely have reported many of these targeted events if targeted surveys with greater depth or higher cadence (e.g., DLT40; Tartaglia et al. 2018) had not discovered them first (Foley et al. 2018). Untargeted surveys independently discovered 94% of the Foley et al. (2018) sample.

Foundation aims to compile a sample of up to 800 SNe Ia observed in *griz* over the next several years. The Foundation Data Release 1 (DR1) includes 180 SNe Ia that pass the Foley et al. (2018) sample criteria for inclusion in a cosmological analysis, approximately equal to the number of published, cosmologically useful SNe Ia from all previous low-$z$ samples combined (Betoule et al. 2014; J18; S18). The Dark Energy Survey cosmological analysis (DES Collaboration et al. 2018), for example, uses just 122 low-$z$ SNe Ia.

Here we combine Foundation data with high-$z$ data observed using the same telescope and the same photometric system, creating for the first time a unified sample with a significant number of SNe Ia ($>25$) at both $z < 0.1$ and $z > 0.1$ observed on the same photometric system. The data reduction and calibration of these data are nearly homogeneous, with the caveats that the use of redder bands at high redshift and the exposure time increase for the high-$z$ data introduce modest sample-to-sample differences. We use both the subset of SNe from the PS1 Medium Deep Survey (MDS) with spectroscopic classifications and the full photometrically classified MDS sample. The photometrically classified MDS sample includes ~5% core-collapse (CC) SN contamination, but this contamination can be marginalized over in a Bayesian framework (Kunz et al. 2007; Hlozek et al. 2012; Jones et al. 2017).

We expand the CC SN simulations in Jones et al. (2017) to gain improved constraints on the effect of CC SN contamination on our cosmological measurements. We also measure star

formation rates (SFRs) for the entire data set using PS1 and SDSS photometry, where available, to test the effect of star formation on SN Ia shape- and color-corrected magnitudes, ensuring that uncertainty in the relationship between SN Ia properties and their host galaxies is not biasing the cosmological parameters.

Because of the limited wavelength coverage of the SALT2.4 model that is used to measure SN Ia distances, the present analysis is restricted primarily to the Foundation *gr* photometry (six Foundation SNe are at high enough redshift to include *i* data). The SALT2 color law is only trained from 2800 to 7000 Å (Guy et al. 2007, 2010; Betoule et al. 2014). Relative to the previous low-$z$ data, this reduces the precision of the measured color of each SN and limits our ability to verify the SALT2 color law with Foundation data. However, an extended SALT2 model, retrained with Foundation data, would allow us to take advantage of the available *iz* observations in the future.

In Section 2, we present the MDS and Foundation data sets. In Section 3, we outline our cosmological parameter estimation methodology. Our results are in Section 4, including a discussion of the consequence of replacing the current low-$z$ sample with the Foundation sample. We discuss future prospects for cosmology with the Foundation data set in Section 5, and our conclusions are in Section 6.

## 2. Data

This paper presents the union of the Foundation and PS1 MDS data sets, both assembled using the PS1 telescope. The extremely well-calibrated PS1 photometric system (Schlafly et al. 2012) makes it ideal for cosmological analyses of SN Ia, which typically have photometric calibration as the dominant systematic uncertainty (e.g., Sullivan et al. 2011; Betoule et al. 2014; Rest et al. 2014; Scolnic et al. 2014a).

The PS1 MDS observed 70 deg$^2$ of sky over approximately 4 yr, with a typical observing sequence of $g_{P1}$ and $r_{P1}$ on the same night, followed by $i_{P1}$ and $z_{P1}$ on the second and third nights, respectively. The $y_{P1}$ filter was primarily used in bright time and does not currently have the depth or precise calibration necessary for a cosmology analysis using high-$z$ SNe. Further details about the MDS strategy are given in Chambers et al. (2016).

The PS1 MDS discovered approximately 5200 SNe, ~350 of which are spectroscopically classified SNe Ia. The spectroscopically classified SNe Ia were used to measure cosmological parameters in S18, Rest et al. (2014), and Scolnic et al. (2014a). Additionally, we measured ~3200 host galaxy redshifts, and 1169 of these SNe with either spectroscopic or light-curve-based classifications were used to measure cosmological parameters (Jones et al. 2017; J18). In the present work, we use both spectroscopically and photometrically classified MDS data to measure cosmological parameters. The light curves and host galaxy spectra for likely SNe Ia are available online at 10.17909/T95Q4X and through the Open Supernova Catalog (Guillochon et al. 2017). The remainder, including likely CC SNe and noisy SN Ia, will be published in future work.

The Foundation Supernova Survey uses the PS1 telescope to follow SNe Ia found by surveys that quickly publish new SN discoveries. To be followed by Foundation, SNe must be within the $3\pi$ footprint ($\delta > -30°$), have $z \lesssim 0.08$, and have Milky Way $E(B - V) \lesssim 0.2$ mag. To minimize peculiar velocity uncertainties, the minimum redshift of Foundation is
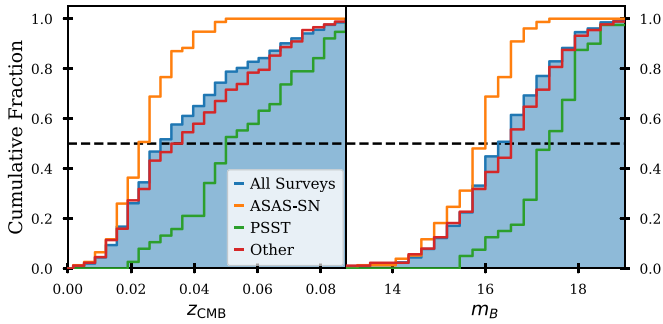
**Figure 1.** Cumulative distributions of the redshift and peak magnitude of SNe Ia in the Foundation DR1 using data from Foley et al. (2018). The two largest SN discovery surveys in this data set are ASAS-SN (orange; 38%) and PSST (green; 20%).

0.015, unless the SNe are near enough to potentially have a Cepheid or tip of the red giant branch distance.

The calibration and sample selection of the Foundation SNe are more similar to the high-$z$ SNe than to previous samples of low-$z$ SNe. The higher-$z$ samples have historically been observed on better-calibrated photometric systems than low-$z$ samples. Foundation, however, uses the PS1 photometric system, which has systematic uncertainties on the few mmag level (Schlafly et al. 2012; Scolnic et al. 2015). Similar to high-$z$ SNe, most Foundation SNe are also discovered by surveying a given area on the sky rather than targeting a preselected set of galaxies, with the majority of the sample coming from ASAS-SN and PSST (Figure 1). In previous samples of low-$z$ SNe, the systematic error due to selection effects was greatly increased by uncertainty over whether the surveys were predominantly magnitude limited or volume limited (J18; S18). A simple yet important advantage of the Foundation Supernova Survey is that we understand that our sample is magnitude limited, which can reduce our final systematic uncertainty due to selection effects by ~50% (Section 4.3).

The first Foundation data release includes 225 SNe Ia. SNe are observed in $griz_{P1}$ at each epoch, with a median cadence of 8 days overall and 5.5 days within 10 days of peak. A total of 180 pass the criteria presented in Foley et al. (2018) for inclusion in a cosmological analysis (see Section 2.3), and 175 of those are included here.[16] Redshifts and classifications for Foundation SNe are given by Foley et al. (2018) and references therein. Foundation will eventually obtain light curves for ~800 SNe Ia to match the *WFIRST* low-$z$ sample requirement (Spergel et al. 2015).

### 2.1. Photometric Pipeline Processing

The photometric pipeline and template construction used to measure the light curves of Foundation SNe is reviewed in Foley et al. (2018) and is in large part based on the `photpipe` pipeline (Rest et al. 2005, 2014). The pipeline for the PS1 MDS is nearly identical and is reviewed in S18. In brief, Foundation image templates were taken from stacked exposures created by the PS1 survey, while PS1 MDS templates were created from the medium deep survey itself after excluding images taken during the season in which a given SN was observed. Both images and templates are resampled and astrometrically aligned to match a skycell in the PS1 sky tessellation. An image

---

[16] Two SNe do not pass the sample cuts when the wavelength range of the SALT2 model is reduced to <7000 Å, and three are below our minimum redshift of $z = 0.015$.

zero-point is determined by comparing point-spread function (PSF) photometry of the stars to updated stellar catalogs of PS1 observations (Chambers et al. 2016). The PS1 templates are convolved to match the nightly images, and the convolved templates are subtracted from the nightly images with HOTPANTS (Becker 2015). Finally, a flux-weighted centroid is found for each SN position, and PSF photometry is performed using "forced photometry": the centroid of the PSF is forced to be at the SN position. The nightly zero-point is applied to the photometry to determine the brightness of the SN for that epoch.

There are multiple systematic uncertainties related to this process, but all of them are on the mmag level. Foundation templates are ~2 mag deeper than the individual exposures and are created from PS1 $3\pi$ survey data. Any expected systematic uncertainties from remaining SN light or the manner in which templates are subtracted (e.g., image subtraction vs. scene modeling) are ~1 mmag (S18). There is an additional 1 mmag systematic uncertainty due to the fact that, for Foundation, forced photometry is performed on the SNe but not on the stars (Figure 2). This systematic uncertainty is relatively low compared to Rest et al. (2014) because >90% of the observations of any given Foundation SN have multiple exposures with signal-to-noise ratios (S/Ns) > 20.

The processing of MDS versus Foundation data has only a few subtle differences (below). Figure 2 shows that the expected bias from these differences is negligible.

1. The photometry of Foundation DR1 light curves was measured using DOPHOT (Schechter et al. 1993), while DAOPHOT (Stetson 1987) was used for the PS1 MDS. The primary difference between the two methods is that DOPHOT uses a Gaussian model to fit the PSF of each image, while DAOPHOT uses a Gaussian with an empirical lookup table for fit residuals. We have verified that the two methods have a median difference of less than 1 mmag (Figure 2, left), and we anticipate updating the photometry with DAOPHOT in the Foundation second data release.

2. In processing the MDS images, `photpipe` used an astrometric alignment algorithm to align the images. The astrometry of the PS1 warp images—single-epoch images that have been resampled and aligned into a skycell of the PS1 sky tessellation—is currently much more accurate than it was during the MDS, and the `photpipe` astrometric alignment stage is no longer necessary.

3. The MDS used forced photometry to compute individual image zero-points, while the Foundation pipeline uses photometry with a floating centroid (unforced photometry). We have verified that using unforced photometry (in conjunction with DOPHOT) does not bias the data (Figure 2, left) and have estimated just a 1 mmag systematic uncertainty due to this effect (D. M. Scolnic et al. 2019, in preparation).

4. Light curves from the MDS have an additional correction for the host galaxy background noise. Photometric uncertainties are increased such that epochs with no SN light have a reduced $\chi^2$ of 1 and a baseline flux value is added such that the weighted average of the flux measured from these epochs is 0. SNe in the MDS are often much fainter than their host galaxies, and a proper accounting for host galaxy background noise can increase photometric uncertainties by up to a factor of ~2–3

3
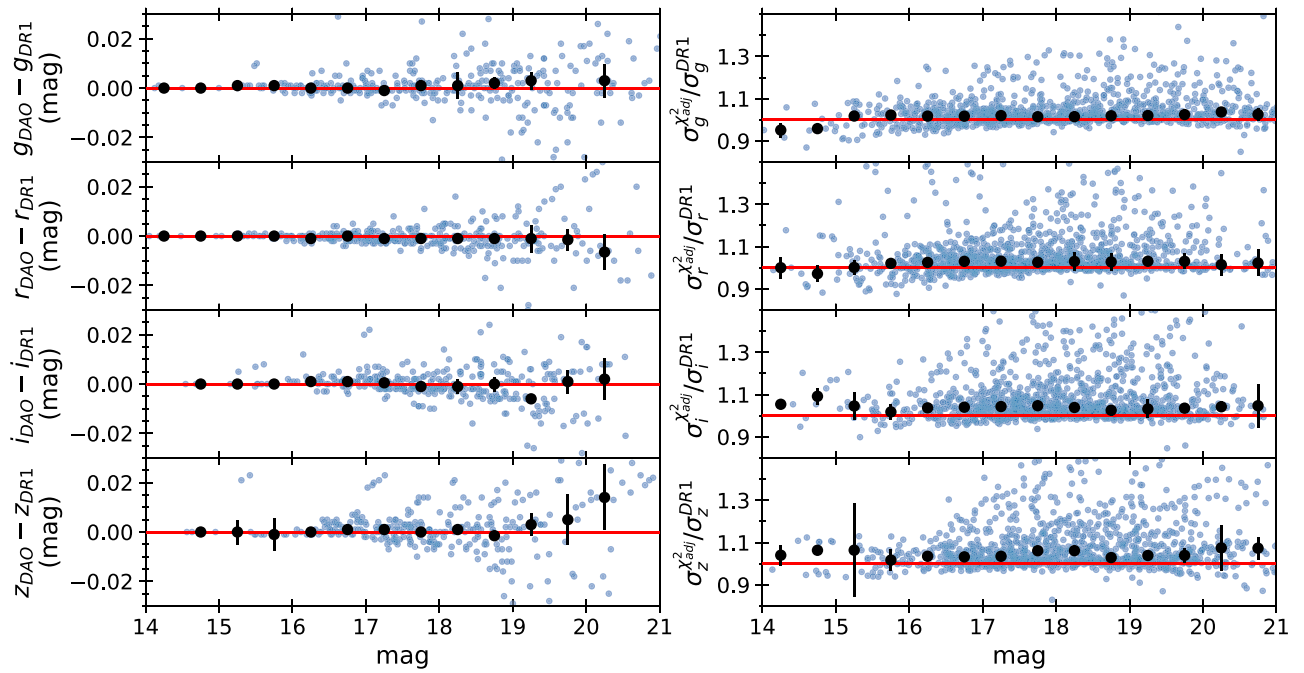
**Figure 2.** Comparison of photometric measurement techniques for the low-$z$ Foundation sample to those used for the high-$z$ PS1 MDS. Left: for a subset of Foundation SNe, the difference between forced DAOPHOT photometry and unforced DOPHOT photometry as a function of magnitude. The median difference is <1 mmag. Right: as a function of magnitude, the fractional increase in uncertainty for Foundation SNe after empirically adjusting the photometric uncertainties using the noise at the SN location in pre-explosion epochs. Performing this procedure on the Foundation DR1 photometry shows that DR1 flux uncertainties are underestimated by just 2%–4% on average.

(Kessler et al. 2015; Jones et al. 2017). However, the effect on Foundation SNe is much smaller.

Host galaxy noise must be added in quadrature to the light-curve flux uncertainties. Because Foundation SNe have much larger fluxes than MDS SNe, they also have larger Poisson uncertainties—a factor of 14 larger within 15 days of peak—and host galaxy noise is therefore a much smaller effect for Foundation. However, we tested the size of the host galaxy noise effect for a subset of the Foundation sample. To do this, we used the fact that each Foundation template is composed of $\gtrsim$10 individual PS1 single-epoch images. Therefore, each SN has $\gtrsim$10 measurements of the background prior to the SN explosion from which the measured uncertainties are compared to the variance of the background. From these measurements, we find that the Foundation DR1 flux uncertainties are underestimated by only 2%–4% on average (Figure 2, right). Because we use PS1 $3\pi$ images as templates for difference imaging, there is correlation between the image flux and the template flux. But, given that there are $\gtrsim$10 warps per template, this correlation is small.

In addition to these subtle differences, the photometry of *both* Foundation and the MDS has been corrected for a subtle bias introduced because the shape of the PS1 PSF (and all PSFs) is dependent on the color of the source. Using the formalism of Guy et al. (2010), this bias can be corrected empirically by a slight linear, wavelength-dependent adjustment to the PS1 $g$-band throughput $T$ when PSF photometry is used:

$$\tilde{T} = T \times [1 + 0.065 \times (\lambda - 4979)/1000]. \qquad (1)$$

The full procedure for improving the color-dependent photometric measurements will be presented in D. M. Scolnic et al. (2019, in preparation), and further details about this bias are given in Guy et al. (2010, see their Figure 4).

### 2.2. Host Galaxy Masses and SFRs

Accurate SN Ia distance measurement requires a correction based on the mass of each SN Ia host galaxy (e.g., Kelly et al. 2010; Lampeitl et al. 2010; Sullivan et al. 2010). There has been some additional evidence that the SFR or specific star formation rate (sSFR) near the site of the SN explosion might correlate with measured SN Ia distances (Rigault et al. 2013, 2015, 2018; Kim et al. 2018; Roman et al. 2018), although these dependencies may be partially included in a correction based on the host galaxy mass. Some of this evidence was also disputed in Jones et al. (2015). However, given that we have the largest current sample of cosmologically useful SNe Ia, it makes sense to simultaneously measure masses and sSFRs to both test for correlations and simultaneously constrain cosmological bias.

To measure host galaxy masses and sSFRs, we measure PS1 $grizy_{P1}$ photometry and SDSS $ugriz$ for each host galaxy. For Foundation, we use PS1 $3\pi$ images from the publicly available first data release (Flewelling et al. 2016). For the PS1 MDS data, however, $3\pi$ images are contaminated by SN light, and therefore we use single-season templates from the PS1 MDS fields as discussed in J18 and S18. The most likely host galaxy for all SNe is determined by using the $R$ parameter method (Sullivan et al. 2006, their Equation (1); see also J18), which uses the proximity, size, and orientation of the galaxy. Nearly all low-$z$ hosts were validated by ensuring that the host redshift matches the SN redshift. For the high-$z$ sample, Jones et al. (2017) estimated that $1.2\% \pm 0.5\%$ of these host galaxy
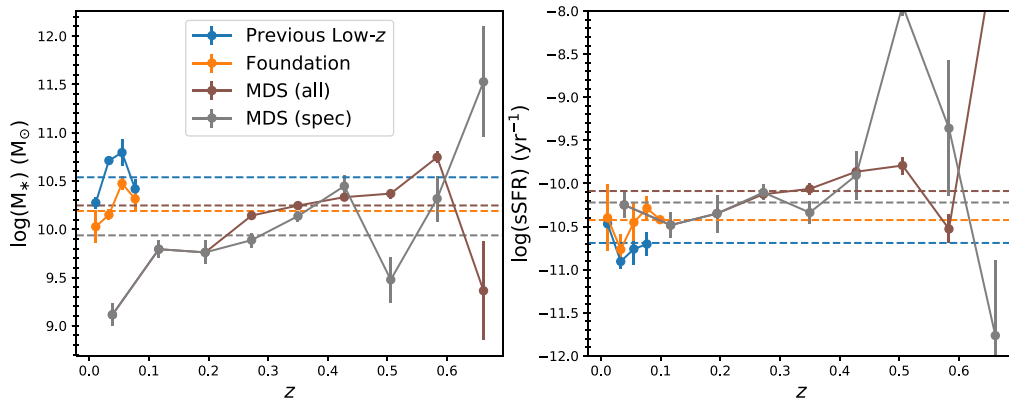
**Figure 3.** Host galaxy mass and sSFR as a function of redshift for the previous low-$z$ sample (blue), the Foundation sample (orange), the full MDS sample (brown), and the spectroscopically classified subset of MDS SNe Ia (gray). Dashed lines indicate the median host mass and sSFR of each sample. We show median bins with uncertainties estimated from the median absolute deviation. Foundation is a closer match to the host galaxy properties of the high-$z$ PS1 data than to those of the low-$z$ sample. In the MDS sample, low-$z$ SNe in faint hosts may have been favored for follow-up as candidate higher-$z$ objects.

determinations were incorrect, an effect that will not significantly bias measuring the correlation of Hubble residuals with host mass and sSFR (see also Gupta et al. 2016).

We measure photometry within a fixed elliptical aperture given by the SExtractor (Bertin & Arnouts 1996) isophotal radius in the PS1 $r$ band. We increase the radius to account for the PSF sizes of each filter and instrument, and we mask out SExtractor-identified contaminating objects in each image. We verify that our measured aperture magnitudes are consistent with the catalog magnitudes provided for each instrument, finding an average offset of <0.05 mag. Some offsets can occur because the catalog magnitudes do not always capture all of the flux from each galaxy and use different effective galaxy radii for different photometric filters.

We use LePHARE (Arnouts & Ilbert 2011) with Bruzual & Charlot (2003) spectral templates to estimate galaxy stellar masses and sSFRs. Galaxy SED templates correspond to spectral types SB, Im, Sd, Sc, Sbc, Sa, S0, and E. $E(B-V)$ is allowed to vary from 0 to 0.4 mag during the fit. We estimate uncertainties by generating Monte Carlo realizations of the host galaxy photometry. For each filter, we generate mock photometric points from a normal distribution with a standard deviation equal to the photometric uncertainties. We then use LePHARE to fit SEDs to each realization of the photometric data, and the uncertainty on the host mass and SFR is given by the spread in output values.

The average host mass and sSFR are shown as a function of redshift in Figure 3 for the Foundation sample, previous low-$z$ samples, and the MDS samples. The Foundation sample is significantly less biased toward high-mass and low-sSFR hosts than previous low-$z$ samples. If SNe Ia depend on their host galaxies in an unexpected way, an analysis with Foundation SNe will not be as systematically affected as one using previous low-$z$ data. See also Figure 8 of Brown et al. (2019) for a comparison between SN Ia host galaxy masses in the targeted LOSS sample (Li et al. 2011) and SN Ia host masses in the volume-limited ASAS-SN sample, as ASAS-SN discovers many SNe that are followed by Foundation. We also note that the full MDS sample selection is less biased than the spectroscopically classified sample, in which faint, low-$z$ galaxies may have been thought to be at higher redshift and therefore targeted for follow-up. The full MDS sample necessarily excludes most apparently hostless SNe, as a redshift cannot be obtained, but the redshift follow-up

discussed in Jones et al. (2017) is otherwise complete to $z \approx 0.3$. At $z > 0.3$, the sample is biased toward brighter, more massive host galaxies for which redshifts can be more easily obtained. The correlation between Hubble residuals and host galaxy mass/sSFR is given for the Foundation sample in Sections 4 and 5.

### 2.3. SALT2 Distances and Selection Criteria

We use the SALT2.4 light-curve fitting method (Guy et al. 2010; Betoule et al. 2014) as implemented in the SuperNova ANAlysis software (SNANA; Kessler et al. 2009) to measure the shape, color, and flux parameters of SNe Ia in this sample. The relationship between SALT2 parameters and the SN Ia distance is given by the Tripp formula (Tripp 1998):

$$\mu = m_B - \mathcal{M} + \alpha \times x_1 - \beta \times c + \Delta_B(z) + \Delta_M, \quad (2)$$

where the light-curve stretch parameter, $x_1$, the color parameter, $c$, and the amplitude parameter, $m_B$, are measured from each SN light curve. Nuisance parameters $\alpha$ and $\beta$ are free parameters that we will estimate simultaneously with SN distance in Section 3. $\mathcal{M}$, a combination of the absolute SN Ia magnitude in the $B$ band at peak and the Hubble constant, is marginalized over during cosmological parameter estimation. $\Delta_B$ is the correction for the redshift-dependent distance bias, which is computed from simulations of our sample (Section 3.1), and $\Delta_M$, the "mass step," is a step function that depends on the host galaxy mass for each SN (see Section 5.3 for alternate parameterizations). SN Ia uncertainties include redshift uncertainty and lensing uncertainty ($\sigma_{\mathrm{lens}} = 0.055z$; Jönsson et al. 2010).

We use these SALT2 parameters to apply the standard sample selection criteria used by J18, which in turn are based on Betoule et al. (2014). These selection criteria include cuts on the SALT2 shape ($-3 < x_1 < 3$) and color ($-0.3 < c < 0.3$) that ensure that the SNe Ia are within the parameter space covered by the SALT2 model training. Cuts on the shape and time of maximum light uncertainty ensure that these parameters are relatively well measured ($\sigma_{t0} < 2$ rest-frame days and $\sigma_{x1} < 1$).

We remove light curves without any epochs between 5 and 45 days after maximum light to avoid multimodal probability distribution functions in the light-curve parameters (Dai & Wang 2016). We also apply a relatively loose cut using the

fit probability, based on the $\chi^2$ and number of degrees of freedom, that the data are consistent with the SALT2 model (*Fitprob* > 0.001). Finally, after fitting with the SALT2 model, we remove up to two photometric outlier points lying >3$\sigma$ from the model fit and then rerun the fitting. Because SALT2 fits a median of 35 light-curve points for MDS SNe, removal of up to two outliers does not greatly affect the fitting. In cases where there are more than two outliers, we remove the two most extreme outlier points. These outliers can be caused by image defects or poor subtractions; however, removing too many epochs could make CC SNe appear more like SNe Ia, which would negatively affect our ability to classify them in the future (Section 3.2). Even when SNe are spectroscopically classified as Type Ia, we only remove a maximum of two outliers for consistency across the sample.

For the Foundation sample, the Foley et al. (2018) cuts also remove spectroscopically peculiar SNe Ia and require the first light-curve point to have a phase of less than +7 days. We require a minimum of eight Foundation light-curve observations in the *gr* bands.[17] The uncertainty on the time of maximum light must be less than 1 day, rather than the looser requirement of 2 days for the MDS sample.

The effects of these cuts on the final sample size are given in J18, Table 1, and Foley et al. (2018), Table 7. For the high-$z$ sample, some of the most significant cuts are the $x_1$ uncertainty cut, which reduces the sample by 27%, and the $x_1$ and $c$ cuts, which each reduce the sample size by ~20%. For the Foundation sample, all SNe are spectroscopically classified as SNe Ia and the S/N of all data are higher. Because of this, only a single SN fails the $x_1$ uncertainty cut, while 6% and 1% fail the $x_1$ and $c$ cuts, respectively.

## 3. Cosmological Parameter Estimation

The steps for estimating cosmological parameters from Foundation and MDS SNe are presented in this section:

1. In Section 3.1, we correct SN Ia light-curve parameters, measured by fitting with SALT2 as discussed above, for sample selection or distance-dependent biases (often referred to as Malmquist biases). Simulations of the SN Ia sample give a prediction for the bias in SN distance measurement as a function of redshift.
2. In Section 3.2, measuring distances from a sample with CC SN contamination requires an estimate of the probability that each SN is a CC SN or Type Ia. These must be estimated from SN light-curve classification (e.g., PSNID; Sako et al. 2011) and are used as priors in the SN likelihood formalism.
3. In Section 3.3, we apply a likelihood function for estimating distances from a sample with both SNe Ia and CC SNe. We estimate distances and uncertainties at a set of 25 log($z$)-spaced redshift "control points" from $0.01 < z < 0.7$.
4. In Section 3.4, for each systematic uncertainty, we repeat step 3 after adjusting the SN light curves for that systematic uncertainty. We generate a combined covariance matrix from all systematic and statistical uncertainties.

5. In Section 4, after including constraints from CMB measurements, BAO measurements, and the local measurement of $H_0$, we estimate final cosmological parameters with CosmoMC (Lewis & Bridle 2002).

We use two different algorithms to correct for selection biases (Section 3.1) and marginalize over CC SNe (Section 3.3). These algorithms are based on the Bayesian Estimation Applied to Multiple Species (BEAMS; Kunz et al. 2007) method of marginalizing over CC SNe, which is discussed in Section 3.3 below. The first is the approach given in J18, which we will refer to as the Photometric Supernovae with BEAMS method (PSBEAMS). The second is the BEAMS with Bias Corrections method (BBC; Kessler & Scolnic 2017). We will discuss each of these methods below and present separate measurements of $w$ using each algorithm. These algorithms differ in the implementation of bias corrections (1D vs. 5D), in the modeling of CC SNe, and in the ways in which events in different redshift bins are combined. The final constraints on $w$ will use the BBC method, as discussed in Section 4.

### 3.1. Simulating the SN Ia Sample and Correcting for Selection Biases

Measurements of SNe Ia in magnitude-limited samples will have distance biases that are caused by SN selection effects. Primarily, these are due to the fact that it becomes more difficult to detect fainter SNe at greater distances. As the S/Ns of the photometric measurements become lower, secondary effects such as biases caused by the $-0.3 < c < 0.3$ and $3 < x_1 < 3$ box cuts come into play.

We predict the distance biases affecting our sample using SNANA simulations. The simulations generate realizations of the SALT2 model after applying real detection efficiencies, survey zero-points, PSF sizes, sky noise, host galaxy noise, and other observables from the real survey. Simulations must be tuned so that the uncertainties as a function of magnitude match the data. One of the primary underlying reasons is that in SNANA simulations, the entire PSF contributes to the noise computation, whereas in the data we must choose a fixed radius within which to fit the PSF and estimate the uncertainties. Simulations also require that the underlying distributions of SN light-curve shapes and colors be robustly determined. These distributions were inferred using the method of Scolnic & Kessler (2016) and will be discussed in more detail in D. M. Scolnic et al. (2019, in preparation).

The SNANA simulations are complex, and detailed discussions of the MDS simulations are given in Jones et al. (2017) and J18. Here we focus on the Foundation simulations, which are illustrated in Figure 4. See also Kessler et al. (2019) for a detailed description of the SNANA simulation process (an overview is given in their Figure 1).

Approximately 60% of the Foundation DR1 data are discovered by two magnitude-limited surveys: ASAS-SN (Holoien et al. 2017) and the PSST (Huber et al. 2015). The spectroscopic follow-up observations of ASAS-SN-discovered SNe are 100% complete (Holoien et al. 2017), while the PSST follow-up observations are incomplete and must be modeled empirically. In addition to biases from the photometric discovery surveys, there are additional complex biases from spectroscopic follow-up. ASAS-SN, for example, is 100% spectroscopically complete, but Foundation can only follow

---

[17] Foley et al. (2018) require 11 total light-curve points in *gri*. However, for the majority of the sample we use only *gr*, so we slightly loosen this requirement.
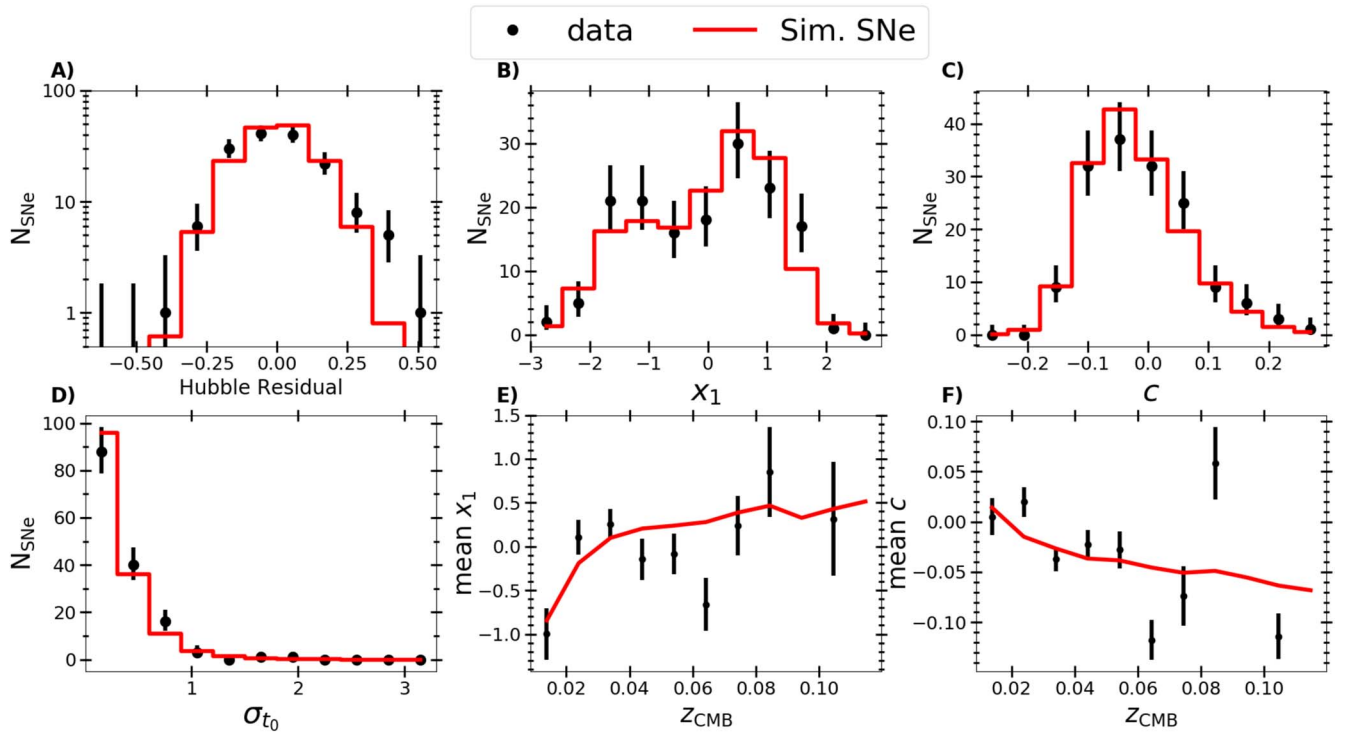
**Figure 4.** Foundation data (black points) compared to G10 SNANA simulations (red). Panels (a)–(d) show Hubble residuals, $x_1$, $c$, and uncertainty in the time of maximum light, respectively. Panels (e) and (f) show the average $x_1$ as a function of redshift and average $c$ as a function of redshift, respectively.

SNe Ia if the spectra are obtained before maximum light. The other surveys have similar complexities.

We must therefore determine the selection efficiency empirically. We first generate a Foundation simulation without any spectroscopic selection cut and compare the observed peak magnitude distribution to that of the data. Dividing the two distributions gives the efficiency of follow-up observations as a function of magnitude. This procedure follows the one used for the Pantheon sample (S18), and additional details about the Foundation spectroscopic efficiency will be given by D. M. Scolnic et al. (2019, in preparation). When the Foundation sample is larger, it will be possible to empirically model the samples from each discovery survey individually, but currently this approach is limited by statistical uncertainty.

Survey of origin likely contributes to some of the dispersion in SN properties seen in Figure 4, but we also see that the distance bias is most strongly affected by the need to spectroscopically classify SNe Ia rather than by the detection limit of the individual surveys. At higher redshift, PSST has a deep detection threshold and is the dominant discovery survey, but the SNe included in Foundation are still preferentially brighter, bluer, and with broader light-curve shapes than the lower-$z$ samples such as ASAS-SN. ASAS-SN SNe are in fact a bit redder and lower-stretch than the average Foundation SN, even though ASAS-SN has a relatively bright detection limit. A Kolmogorov–Smirnov test finds $p$-values of 0.30 and 0.32 for the ASAS-SN $x_1$ and $c$ distributions, respectively, compared to the rest of the Foundation sample, showing that these distributions are consistent with those of Foundation as a whole.

Additionally, ASAS-SN is more efficient at finding SNe near the cores of their host galaxies than other surveys (Holoien et al. 2017). There has been some investigation of whether SNe near their host galaxy centers have biased distance measurements (Hill et al. 2018), but no evidence has yet been found. We explore the

possibility of this bias using the public data of Jones et al. (2018a), which include much of the Foundation sample; we use the SN/host galaxy $R$ parameter to divide the sample into low-$R$ and high-$R$ subsets (see the discussion of $R$ in Section 2.2). We find that the measured Hubble residuals are largely insensitive to $R$, with a maximum $1.6\sigma$ difference of $0.058 \pm 0.035$ mag between SNe Ia with $R > 0.5$ and SNe Ia with $R < 0.5$ (likely even less significant than $1.6\sigma$, as we have not accounted for the look-elsewhere effect). Because 14% of the SNe Ia in the high-$z$ sample have $R < 0.5$ and just 7% of the SNe in the low-$z$ sample have $R < 0.5$, we also do not expect a strong redshift-dependent bias even if this effect is measured to be significant in the future. Naïvely, the $z$-dependent bias would be the difference in percent of $R < 0.5$ SNe between low-$z$ and high-$z$ samples multiplied by the size of the difference, which gives a bias of $\sim$3 mmag. We do not model this aspect of possible distance measurement bias, whether due to statistical fluctuation, SN physics, host galaxy dust, or photometric measurement bias, but note that this may be a necessary area to explore as measurements of $w$ with SNe Ia become increasingly precise.

We simulate two models for the scatter of SNe Ia, with distributions of $x_1$ and $c$ for each model given by Scolnic & Kessler (2016). The standard SALT2 error model (Guy et al. 2010, hereafter G10) attributes $\sim$75% of SN dispersion to variation in SN luminosity that is uncorrelated with color. Approximately 25% of the dispersion is given by wavelength-dependent flux variation that is uncorrelated with luminosity. The Chotard et al. (2011, hereafter C11) model, on the other hand, attributes most dispersion to variation in color that is uncorrelated with luminosity. Kessler et al. (2013) translated the G10 and C11 models of broadband covariance into wavelength-dependent models that can be simulated as spectral variations, and these models are the basis for our C11 simulations. We simulate both models, as the data are unable

to distinguish between them. The most significant consequence of changing between these two models is that the inferred $\beta$ changes by $\sim 0.5$ (noted by Scolnic et al. 2014b).

Figure 4 compares the Foundation simulations to the data. In panel (a), there is a slight discrepancy on the faint side of the Hubble residuals; the data show five SNe in this bin, while the simulations predict just one. However, the data in the fainter bins have somewhat larger uncertainties in the $x$-direction, which makes the number of SNe in the fainter bins more subject to statistical fluctuation.

In panel (c), we note that simulating just a single value of $\beta$ appears to reproduce the observed data well. However, there is substantial evidence from UV to near-IR observations that SN Ia extinction laws vary significantly from SN to SN (e.g., Amanullah et al. 2015), which could be partly explained by SN radiation pressure impacting the grain sizes of the dust distribution near the SN (Bulla et al. 2018). There is also evidence that the $\beta$ parameter similarly varies in different types of host galaxies (e.g., Jones et al. 2018a). It is therefore likely that some of the scatter of SNe Ia about the Hubble diagram is due to variation of dust properties. Although we do not fit for $\beta$ for individual SNe (e.g., Burns et al. 2014, 2018), for measurements of $w$ we are concerned primarily with whether the average value of $\beta$—due to the redshift dependence of dust properties or SN Ia properties—is changing with redshift. In Section 4.3, we therefore test whether allowing $\beta$ to evolve with redshift could be systematically biasing our cosmological results. Because SN Ia host galaxies in Foundation are more similar to high-$z$ host galaxies as discussed in Section 2.2, we anticipate that this analysis is less sensitive to this type of potential bias than previous analyses.

Finally, in panels (e) and (f) there are occasionally large bin-to-bin jumps in average $x_1$ and $c$ as a function of redshift. While some of these jumps may be due to statistical fluctuation, it is also likely that our simulations are not a perfect description of the underlying data owing to approximations in modeling the multiple subsurveys that compose Foundation. To explore whether changing the $z$-dependence of $x_1$ and $c$ affects the predicted distance biases, we simulated a Foundation-like survey with an extremely $z$-dependent mean $x_1$ and $c$. We smoothly evolve the mean $x_1$ from $-1$ at $z = 0$ to $+2$ at $z = 0.08$ and $c$ from a mean of 0 at $z = 0$ to a mean of $-0.15$ at $z = 0.1$. Even these extreme variations in the average color result in distance biases that change by a maximum of 3 mmag with no significant systematic trend.

Figure 5 demonstrates that the S/Ns of SNe in Foundation simulations are a better match to the real data than the previous low-$z$ sample simulations were to the previous CfA/CSP data (Riess et al. 1999; Jha et al. 2006; Hicken et al. 2009a, 2009b, 2012; Contreras et al. 2010; Folatelli et al. 2010; Stritzinger et al. 2011). In particular, the highest-S/N SNe from CfA/CSP are not represented in the simulations, and the lowest-S/N data do not precisely follow the prediction given by the simulations. The reduced $\chi^2$ of the CfA/CSP data compared to the simulations is 5.4 (although the exact value depends on the bin sizes). The large reduced $\chi^2$ is largely because of the heterogeneous nature of the low-$z$ SN compilation, which came from a large number of separate surveys, and the sample selection criteria of those surveys are varied. The Foundation data agree more closely with the Foundation simulations, with a reduced $\chi^2$ of 1.9 when compared to the data.
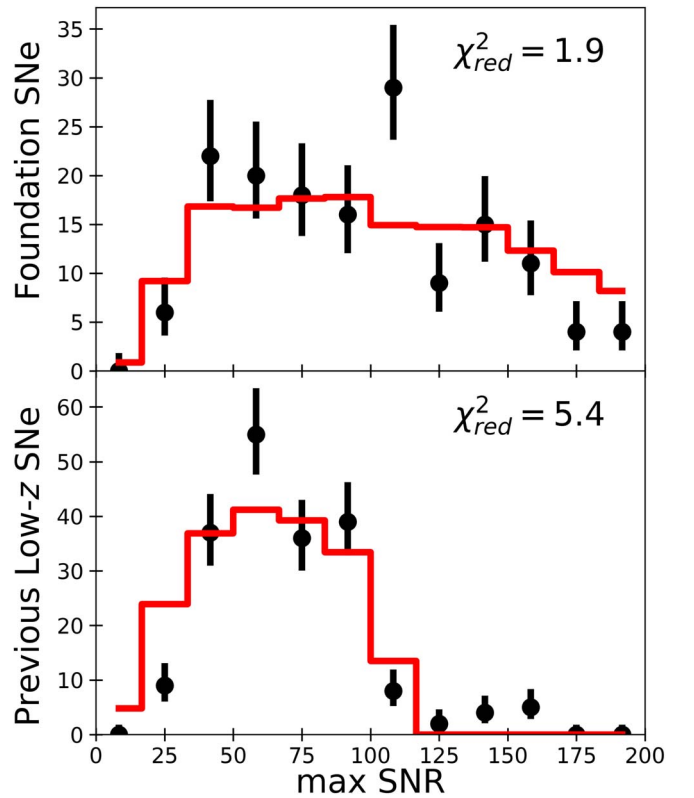


**Figure 5.** Number of SNe in the Foundation sample (top panel) and previous low-$z$ sample (bottom panel) as a function of the peak-brightness S/N (black points), with Poisson uncertainties displayed for each S/N bin. In red, we display the best-fit G10 simulation for each sample. Comparing the data to the simulations, the reduced $\chi^2$ for the Foundation and previous low-$z$ samples are 1.9 and 5.4, respectively.

These simulations are used to predict the distance bias as a function of redshift for each SN sample. The PSBEAMS method of correcting for distance biases, the standard approach prior to the Pantheon analysis (e.g., Conley et al. 2011; Betoule et al. 2014), used SNANA simulations to generate a one-dimensional correction to the SN distance as a function of redshift. An alternative approach is the BBC method (Kessler & Scolnic 2017), which was used to derive cosmological parameter measurements from the Pantheon sample (S18). BBC applies bias corrections to three parameters—$m_B$, $x_1$, $c$— and those bias corrections are also treated as a function of nuisance parameters $\alpha$ and $\beta$. The $\alpha$ and $\beta$ dependences of the bias corrections are included in the BBC likelihood.

The BBC approach reduces the scatter about the Hubble diagram while explicitly correcting the known dependence of Hubble residuals on $x_1$ and $c$. The BBC method is also used to marginalize over CC SNe in the present analysis, but BBC improves the precision of cosmological parameter measurements even when used in an analysis restricted only to spectroscopically classified SNe Ia such as Pantheon (S18). We use both the PSBEAMS method and the BBC method in this analysis.

### 3.2. Simulating the CC SN Sample and Generating Prior Probabilities

Bayesian likelihood models such as those used in this work require accurate simulations of the CC SN contamination in order to validate the method and yield estimates of the prior

probability that each SN is of Type Ia. In previous work, we estimated prior probabilities from one of four different methods. The most reliable of these methods was the PSNID classifier (Sako et al. 2011), which compares noise-free simulations of CC SNe and SNe Ia to the observed photometric data and gives a Bayesian probability that each SN is of Type Ia. The second-most reliable method was Nearest Neighbor (NN; Kessler & Scolnic 2017; Sako et al. 2018), which compares the redshift and the SALT2 $x_1$ and $c$ parameters measured for each SN to simulated distributions of these parameters. NN gives a distance metric from data to simulated SNe Ia and CC SNe to determine the classification. The distance metric relies on three free parameters, one each for $z$, $x_1$, and $c$, which are determined during a training stage.

In J18 we also employed two classifiers that were independent of a number of biases in CC SN simulations (Jones et al. 2017). The first, which we call *Fitprob*, used the $\chi^2$ and degrees of freedom of the SALT2 model fit to estimate the probability that the SALT2 model matches the data. The second, GalSNID (Foley & Mandel 2013; Jones et al. 2017), uses the fact that CC SNe are not found in old stellar populations to estimate the SN Ia probability from host galaxy spectroscopy and imaging alone.

Using simulations, we found that GalSNID and *Fitprob* could bias $w$ by ~0.02, but we used them as systematic uncertainty variants owing to their independence of CC SN simulations. However, these known biases are larger than necessary for the present analysis, and therefore we instead use simulations of different CC SN distributions to predict the small distance biases caused by imperfections in our CC SN classification methods (Section 3.2).

These methods were validated in Jones et al. (2017) but require separate validation for each sample and analysis, as subtle differences in sample selection can change the effectiveness of our method of marginalizing over CC SNe. These validation tests will be used as part of our systematic uncertainty budget (Section 3.4).

We design three SNANA simulations to encapsulate the uncertainty in CC SN rates, luminosity functions, and the representativity of the available templates. We note that these simulations have already been tuned to match the MDS data by Jones et al. (2017), who found that the bright tail of the CC SN distribution is poorly constrained by Li et al. (2011), requiring the peak of the CC SN LFs to be adjusted by ~1 mag.[18] In the first simulation, we remove 50% of the CC SN templates from the simulation to evaluate the effect of having an incomplete CC SN template set. We randomly removed exactly 50% of the templates for each CC SN subtype, so that all subtypes were still represented in the simulations. In the second simulation, we increase the peak CC SN luminosity functions by an additional 0.5 mag, which in turn makes them a higher fraction of the total SN population. In the third simulation, we apply a strong $A_V$ distribution to the data to account for the fact that templates may be preferentially unreddened compared to the real data. These adjusted CC SN models are shown in Figure 6 along with the reduced $\chi^2$ of each model compared to the data. Even though these adjustments are relatively drastic, all CC SN

simulations have a reduced $\chi^2 < 2.5$ and therefore represent conservatively large adjustments to the CC SN simulation parameters that conservatively account for the uncertainties in modeling CC SN contamination.

We then validate the method by replacing our photometrically classified SNe with each simulation, keeping the real spectroscopically classified SNe. We use PSBEAMS and BBC to measure distances from the sample (see below) both with and without including simulated CC SNe. The difference in these two sets of distances is the distance biases that are introduced by our method of marginalizing over the CC SNe. We correct for the distance biases predicted for the baseline simulation and treat the biases from each of the three simulation variants, relative to the baseline biases, as systematic uncertainties. These systematic uncertainties are typically less than 5 mmag and are shown in Figure 7. We note that the BBC method gives large biases at high $z$ for some simulations, but those biases primarily affect some of the largest-uncertainty bins. The PSBEAMS points at $z = 0.7$ also have large uncertainties (as no SNe at $z > 0.7$ are used), and bias in these bins has minimal impact on cosmological parameter estimation. These offsets are included in the systematic uncertainty budget.

### 3.3. The Likelihood Model

Both frameworks that we use to analyze our sample (PSBEAMS and BBC) are based on the BEAMS method, first presented by Kunz et al. (2007) and refined by Hlozek et al. (2012). BEAMS is a Bayesian method for simultaneously modeling multiple "species" that are partially overlapping in some parameter space. In this case, BEAMS models SNe Ia and CC SNe, which overlap on the Hubble diagram. A brief overview of the method is given in the Appendix, with the full mathematical formalism for the PSBEAMS and BBC methods given in J18 and Kessler & Scolnic (2017), respectively.

The primary difference between the methods, apart from the bias correction differences discussed above, is that BBC uses SNANA simulations of the CC SN distribution to put a prior on the $z$-dependent Hubble residuals expected from CC SN contamination. This avoids the assumption of a redshift-dependent functional form for the CC SN distribution. However, a parametric CC SN model from Hlozek et al. (2012) can also be specified within BBC, which allows more flexibility. In contrast to the PSBEAMS method, which uses a parameterized CC SN model that is linearly interpolated in $\log(z)$ space, Hlozek et al. (2012) treat the CC SN distances and dispersion as polynomial functions of $z$.

For the PSBEAMS method we remove likely CC SNe ($P(\mathrm{Ia}) < 0.5$), following J18, finding that the results are more robust and less affected by systematic uncertainty variants after this cut is applied. This cut reduces the sample to 1085 SNe but does not significantly affect the precision of the cosmological constraints. For the BBC method we include likely CC SNe, as our preference is to use all available data, even those data with low weight, when possible.

In addition, unlike the PSBEAMS method, the BBC method does not have a parameter to linearly shift $P(\mathrm{Ia})$ values to adjust the prior probabilities that an SN is Type Ia (Equation (5)). This parameter helps to correct inaccurate light-curve classifications. Because of this difference, for PSBEAMS we use PSNID as the baseline classifier for determining prior probabilities that a given SN is of Type Ia and NN as the baseline classifier for

---

[18] Jones et al. (2017) do not suggest that the peak of the CC SN LFs as measured by Li et al. (2011) is incorrect by ~1 mag, but rather that an ad hoc procedure to brighten LFs is capable of reproducing the observed bright tail of the CC SN distribution. The MDS survey preferentially detected bright CC SNe.
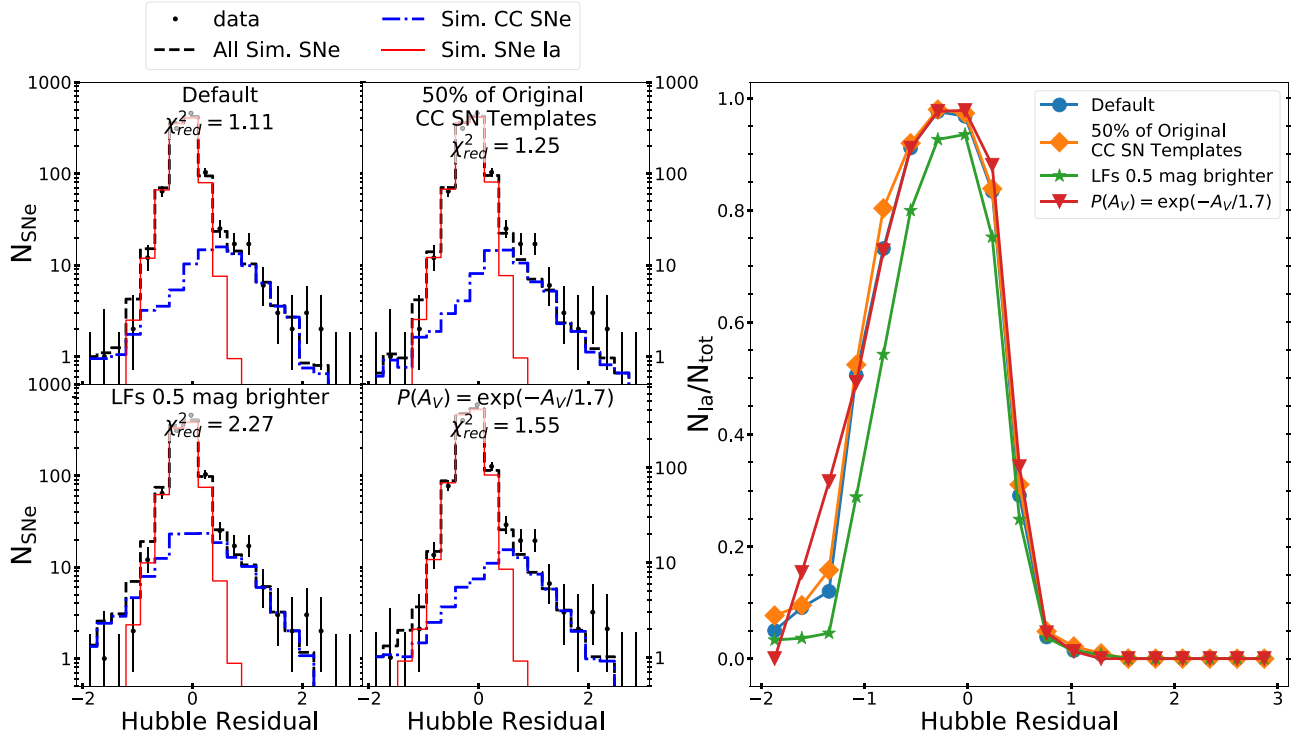
**Figure 6.** Left: adjusted CC SN simulations (blue) compared to the data (black points). Simulated SNe Ia are in red, and the cumulative simulated CC+Ia distribution is in black. From top left to bottom right, we show the default simulations, simulations that reduce by 50% the number of CC SN templates used to simulate the sample, simulations that brighten the intrinsic luminosity of the simulated CC SNe by 0.5 mag, and simulations that add additional extinction to the CC SN templates. Right: fraction of SNe Ia as a function of Hubble residual for each simulation.
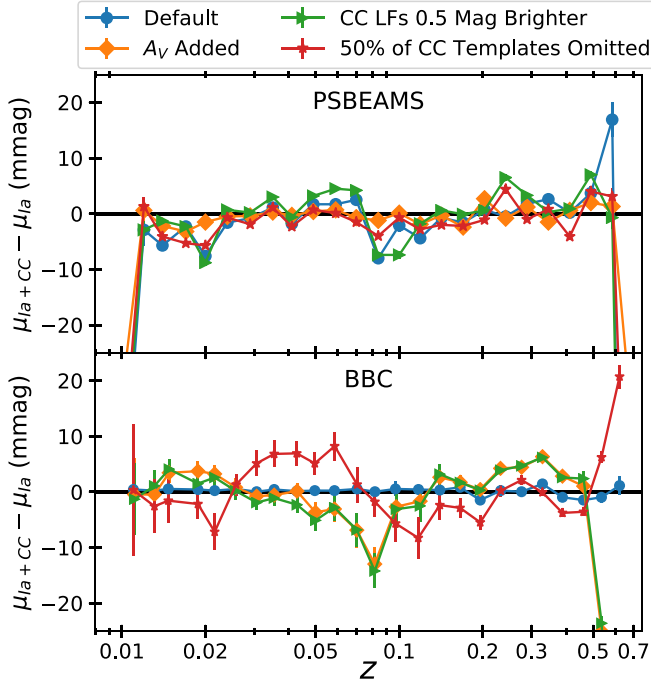


**Figure 7.** For the PSBEAMS method (top) and BBC method (bottom), bias in binned SN Ia distance measurements due to CC SN contamination as a function of redshift. Distances are measured from the real spectroscopically classified sample combined with a simulated photometric (CC SN-contaminated) sample. The simulated photometric sample uses one of four different CC SN simulations, which are discussed in Section 3.4. Biases are typically $\lesssim$5 mmag for all simulated CC SN populations.

BBC, as we found that excluding a linear shift parameter can give biased results when using PSNID probabilities. See the Appendix for further explanation.

### 3.4. Systematic Uncertainties

The systematic uncertainties in this analysis are largely unchanged from the MDS analysis of J18. Therefore, we summarize them here and direct the reader to Section 4 of J18 for a more detailed description of each systematic uncertainty.

The systematic uncertainties affecting SNe in this sample can be attributed to seven broad categories: Milky Way extinction, distance bias/selection effect correction, photometric calibration, SALT2 model calibration, sample contamination by CC SNe or incorrect redshifts, low-$z$ peculiar velocity corrections, and the dependence of SN Ia luminosities on their host galaxies. Additionally, we check for the redshift dependence of $\Delta_M$ and $\beta$ but find no significant evolution in their values. An example from each type of systematic uncertainty is shown in Figure 8, using both the PSBEAMS method and the BBC method.

Replacing the previous low-$z$ sample with the Foundation sample has reduced the low-$z$ distance bias systematic uncertainty from ~1% to 1.5% to just a few mmag. The reduced distance bias systematic uncertainty can be attributed to two effects, both shown in Figure 9. First, S18 and J18 corrected for distance biases considering that $z < 0.1$ SNe may be from either a volume-limited or a magnitude-limited sample, as the data come from surveys that often targeted nearby, massive galaxies. The simulations from J18 and the Pantheon
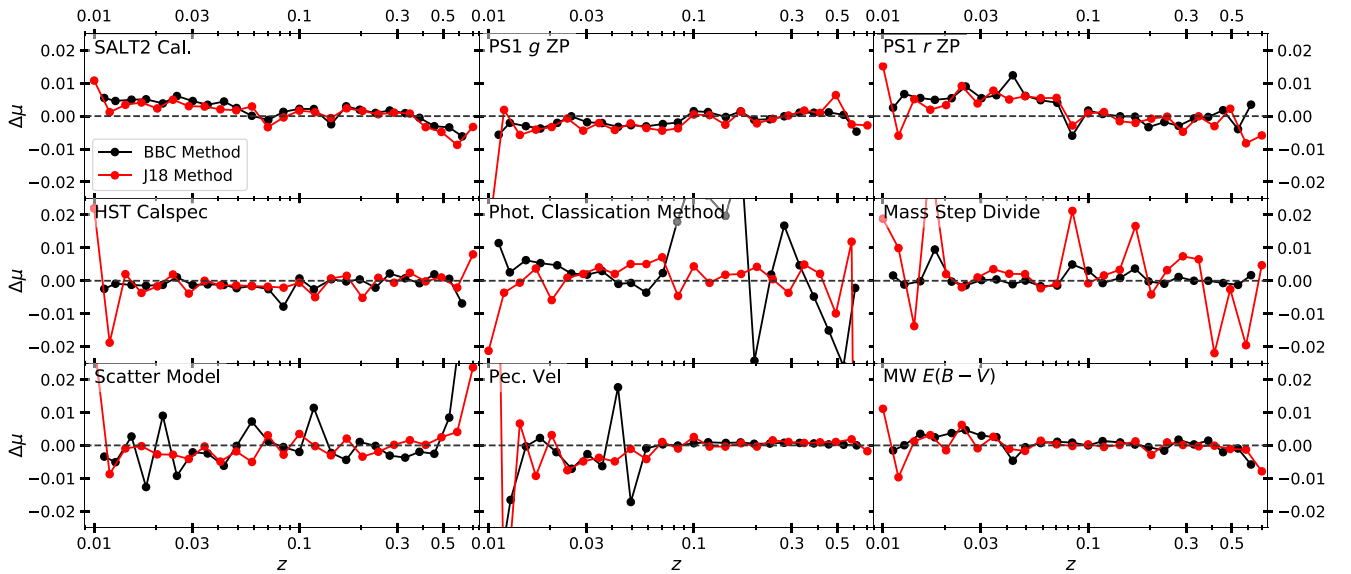
**Figure 8.** Comparison of BBC systematic uncertainties (black) to systematic uncertainties from the PSBEAMS method used in J18 (red). For several of the largest sources of systematic uncertainty, we show how the distances are affected as a function of redshift by that systematic uncertainty (we show a $1\sigma$ shift). "SALT2 Cal.," "PS1 $g$ ZP," "PS1 $r$ ZP," and "*HST* Calspec" refer to the SALT2 calibration uncertainty, PS1 $g$- and $r$-band zero-point uncertainties, and calibration uncertainty in the Calspec system as defined through *HST* observations, respectively. "Phot. classification method" is the shift in distances caused by using a different method of estimating the prior probabilities that each SN is of Type Ia, and "Mass Step Divide" refers to the systematic uncertainty if the divide between what is defined as low- and high-mass host galaxies is shifted by 0.15 dex (within current observational constraints). "Scatter Model" is the systematic uncertainty caused by the difference between the G10/C11 models. "Pec. Vel." and MW $E(B - V)$ are uncertainties in the peculiar velocity correction and the Milky Way extinction.
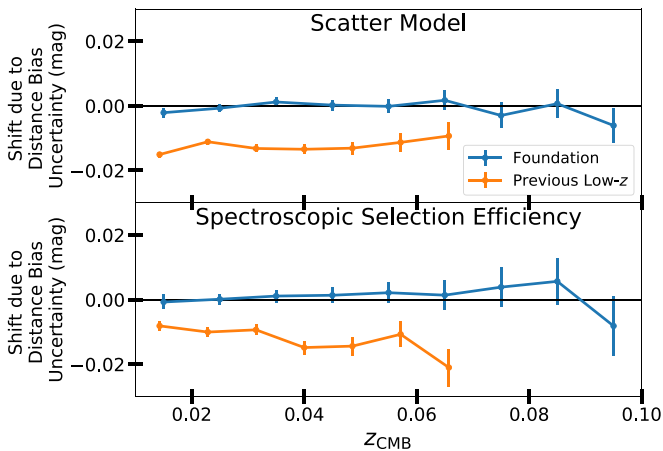


**Figure 9.** Comparison of systematic uncertainties due to bias correction as a function of redshift between the Foundation sample (blue) and the previous low-$z$ SN sample (orange). We show the shift in distance due to the difference between the distance bias predictions of the G10 and C11 models (top) and due to adjusting the uncertain spectroscopic selection efficiency (bottom).

analysis (S18) indicated that even some $z \approx 0.03$ SNe may have been missed if the survey was magnitude limited. For the Foundation Supernova Survey, we understand that our sample is dominated by magnitude-limited data, and therefore we do not include a variant that bias-corrects the sample as though it were volume limited. Second, likely because the Foundation Supernova Survey is not a targeted search, we have a sample with a bluer distribution of SN colors that more closely matches the high-$z$ sample. Foundation SNe have a median $c$ parameter of $-0.020$, while SNe in the previous low-$z$ sample have a median $c$ of 0.004. The high-$z$ PS1 data have a median $c$ of $-0.017$. Comparisons of $x_1$ and $c$ for the Foundation sample compared to previous low-$z$ data are shown in Foley et al. (2018), their Figure 7. Because the average SN color varies less

with redshift, the difference between the distance bias predictions from the G10 and C11 scatter models is much smaller.

The size of the photometric calibration systematic uncertainty has increased, which is due to the fact that we have only a single photometric system in this analysis. Although PS1 calibration uncertainties are low—just 3 mmag per filter (D. M. Scolnic et al. 2019, in preparation; Scolnic et al. 2015)—multiple uncorrelated photometric systems would reduce the calibration systematic uncertainty further, as would the ability to include the Foundation $iz$ observations. In addition, because the bluest band in the sample is $g$, high-$z$ SN observations measure much bluer rest-frame wavelengths than the low-$z$ observations. This increases the impact of the PS1 calibration systematic uncertainties, the SALT2 calibration systematic uncertainties, and the 0.5% slope uncertainty in the *Hubble Space Telescope* CALSPEC system (Bohlin 2014). For context, a 3 mmag change in distance modulus from the median redshift of the Foundation sample ($\sim$0.035) to the median redshift of the MDS sample ($\sim$0.35) would shift $w$ by approximately 1%; however, a shift in a single filter can also bias SN color measurements, and therefore the shift in distance can be larger than 3 mmag in practice (Figure 8, top right panel). The systematic uncertainty associated with the Supercal correction is not included in this analysis (and was included in J18) because all of our data are already on the PS1 photometric system.

We have slightly altered the J18 method of accounting for the systematic uncertainty caused by marginalizing over CC SNe. We first use an alternate method of estimating prior probabilities (either NN or PSNID, depending on the baseline choice for each analysis framework) and treat the change in distances as a systematic uncertainty. Second, for the PSBEAMS framework, we use two alternate models of the CC SN distribution, a skewed Gaussian or two-Gaussian parameterization of the CC SN distribution, and treat the

average of these two parameterizations relative to the baseline parameterization as a systematic uncertainty. When using the BBC framework, we use a simulated CC SN distribution as the baseline CC SN distribution and use the Hlozek et al. (2012) parameterized CC SN model as a systematic uncertainty variant. Additionally, using the four CC SN simulations discussed above, we correct for the distance biases predicted for the baseline simulation and treat the biases from each of the three simulation variants, relative to the baseline biases, as systematic uncertainties. Finally, we test the effect of fixing nuisance parameters to the values determined from just the spectroscopically classified data and include the differences in distance in our systematic uncertainty budget.

Incorrect redshifts due to noisy spectra and mismatched host galaxies are also a source of contamination that is discussed in detail in Jones et al. (2017). In the high-$z$ sample presented here, we expect of order $1.2\% \pm 0.5\%$ of the redshifts to be incorrect. We implicitly model these as part of the CC SN distribution, as simulations show that they are nearly always outliers on the Hubble diagram and do not follow the same distribution as SNe Ia. As discussed in Jones et al. (2017), simulations show that these outliers are not a major source of systematic error. See Roberts et al. (2017) for an alternative method of accounting for redshift contamination that includes incorrect redshifts in the BEAMS likelihood itself.

Finally, extinction by dust in the IGM is a potential source of systematic uncertainty (Ménard et al. 2010; Goobar et al. 2018). Goobar et al. (2018) argue that a slight discrepancy in the low-$z$ data used by the Pantheon sample could be due to IGM dimming. However, we note that the Foundation sample is slightly fainter than $\Lambda$CDM (Section 4), whereas the Pantheon low-$z$ data were slightly brighter, and therefore it appears more likely that the differences in distance between Pantheon and $\Lambda$CDM at low $z$ are due to survey-specific systematic uncertainties. We do not include this effect in our systematic uncertainty budget, as it is not detected in the Foundation sample, but note that future Foundation analyses, when the $iz$ bands are included in the analysis, will be better able to put constraints on the ways in which IGM dust and variable $R_V$ affect SN Ia distances.

## 4. Results

In this section, we begin by discussing the change in distances when the previous sample of low-$z$ SNe Ia (CfA/CSP SNe) are replaced by Foundation SNe Ia. Distances and cosmological parameters for the combination of CfA, CSP, and MDS SNe were reported in J18; therefore, we begin with the PSBEAMS method from J18 of correcting for distance biases and marginalizing over CC SNe to allow a direct comparison to the J18 results. We then discuss the results from the alternate BBC method, which reduces the dispersion of the SN Ia sample and therefore improves the precision of the results, albeit with additional differences in methodology.

The full SN Ia sample used to measure cosmological parameters from the Pan-STARRS1 telescope is shown in Figure 10. The binned distance residuals from SNe Ia (after marginalizing over all nuisance parameters) are shown relative to the binned distance residuals from the previous low-$z$ SN sample in Figure 11. For a given redshift, Foundation SNe Ia have larger distances on average, and a corresponding positive shift in Hubble residuals, compared to SNe Ia from the previous low-$z$ sample. The change in distances from the

$z < 0.1$ sample is $0.046 \pm 0.027$ mag (a weighted average including systematic uncertainties and covariances), a significance of $1.7\sigma$. We note that this difference is assuming the average of the C11 and G10 scatter models. If only the G10 scatter model is used (neglecting the scatter model systematic uncertainty), the difference would be $0.030 \pm 0.023$ mag, which has a lower significance of $1.3\sigma$. Similarly, using only the C11 model gives a $0.062 \pm 0.023$ mag difference, which may suggest that the G10 model is favored by these data. Systematic and statistical uncertainties on this difference are approximately equal. We note that the value of $\Delta_M$ measured by PSBEAMS in this work ($0.088 \pm 0.013$) is 0.014 mag lower than than the $\Delta_M$ value measured in J18, too small of a shift to cause the observed shift in distances.

To test whether these low-$z$ distances could be systematically affected by fitting light curves where the bluest band is rest-frame $g$, we refit the CSP data with $gr$ photometry alone (neglecting $BV$ data), finding that distances were an average of 1.5% fainter at $2.8\sigma$ significance from statistical uncertainties alone. However, this possible bias is just one-third of the total shift in SN Ia distance when the previous low-$z$ data are replaced by Foundation. Furthermore, as SNe Ia are better calibrated and better standardizable candles at redder wavelengths, it may be that the $gr$-only results are less subjected to systematic SALT2 training or calibration uncertainties than the CSP $B$ measurements.

### 4.1. Nuisance Parameters

Nuisance parameters from the PSBEAMS and BBC analysis frameworks are reported in Table 1. All PSBEAMS measurements are consistent with the measurements from previous low-$z$ data combined with MDS (reported in J18), although the systematic uncertainties due to the intrinsic dispersion are significantly lower in the present analysis. This may be because Foundation SNe Ia have a lower dispersion that is more similar to the MDS data (Foley et al. 2018).

The SALT2 $\alpha$ and $\beta$ parameters from the BBC method are slightly different from those from the PSBEAMS method. With BBC we find $\alpha = 0.143 \pm 0.005$ and $\beta = 3.218 \pm 0.071$ (stat. error only); compared to the PSBEAMS method, $\alpha$ is lower by 0.019 and $\beta$ is higher by 0.092. The difference in $\alpha$ is marginally statistically significant and must be driven by the Foundation sample, as both J18 and S18 measured the value of $\alpha$ from the MDS sample to be $>0.16$. With the BBC method, we measure $\alpha$ from the Foundation sample alone to be $0.137 \pm 0.013$.

The reason why the Foundation sample $\alpha$ would be particularly affected by observational biases is unclear, and such a bias is not recovered in simulations of the sample. From simulations, we confirmed that both PSBEAMS and BBC recover $\alpha$ and $\beta$ accurately in a spectroscopically classified sample, though when CC SN contamination is included, $\beta$ is biased by an average of $-0.03$ by the BBC method and by $+0.09$ by the PSBEAMS method (see Jones et al. 2017 for more discussion of these biases). Similarly to the Pantheon results, BBC reduces the sample dispersion substantially—by 16% in this analysis—compared to 1D bias correction methods.

The mass step, $\Delta_M$, from BBC is $0.044 \pm 0.010$ mag, lower than the PSBEAMS mass step by 0.044 mag. We find that the difference between the BBC and PSBEAMS mass step is driven by the strong implicit dependence of the average $x_1$ bias correction on host galaxy mass. The photometric sample has a
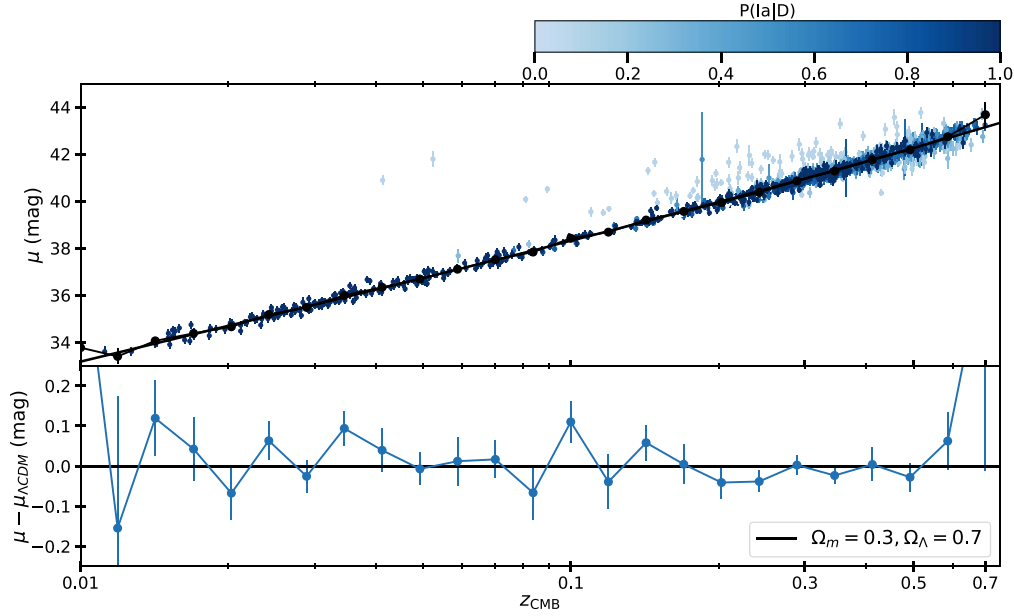
**Figure 10.** Hubble diagram (top) and Hubble residuals (bottom) from the combined Foundation and MDS sample. In the top panel, opacity is set using the approximate posterior probabilities, $P(\text{Ia}|D)$, for the photometrically classified data. In the bottom panel, the points and the lines connecting the points represent the piecewise-linear function of $\log(z)$ that we use to fit the SN Ia distances (see the Appendix). Note that the highest- and lowest-redshift control points have extremely high uncertainties, as no SNe are above or below them in redshift, respectively. Residuals are shown compared to a nominal flat $\Lambda$CDM model with $\Omega_m = 0.3$ and $\Omega_\Lambda = 0.7$.
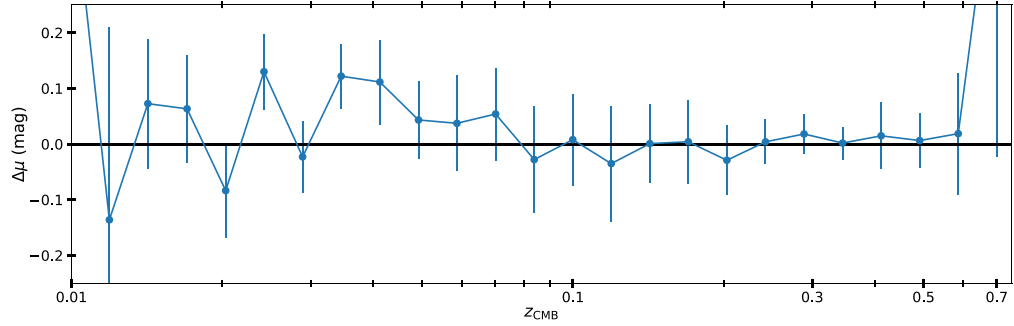


**Figure 11.** As a function of redshift, Hubble residuals from the combined Foundation+MDS sample subtracted by those from the previous low-$z$+MDS sample (J18). Foundation distances are $0.046 \pm 0.027$ mag fainter than those from the previous low-$z$ sample, which gives a positive change in measured $w$. The highest- and lowest-redshift points have extremely high uncertainties, as no SNe are above or below them in redshift.

number of high-mass host galaxies, as these are preferentially more likely to yield redshifts, and SNe Ia in these galaxies have narrower light curves on average than those in low-mass galaxies (Howell 2001). SNe with narrower shape parameters tend to have somewhat negative Hubble residuals owing to observational bias (Scolnic & Kessler 2016), an effect corrected by the BBC framework and not PSBEAMS. A lower mass step from the BBC method was also seen in the Pantheon analysis. These findings also agree with Brout et al. (2019), who see a positive correlation between the size of $\Delta_M$ measured from a given sample and the intrinsic dispersion of that sample.

### 4.1.1. Nuisance Parameter Evolution

We test for redshift dependence of $\beta$ and $\Delta_M$ by allowing those parameters to evolve linearly with redshift:

$$\Delta_M = \Delta_{M,0} + \Delta_{M,1} \times z,$$
$$\beta = \beta_0 + \beta_1 \times z. \tag{3}$$

**Table 1**
Dependence of Nuisance Parameters on Analysis Method for MDS +Foundation

| | PSBEAMS Method | | BBC Method | |
|---|---|---|---|---|
| | | $\sigma_{\text{stat}}$ | | $\sigma_{\text{stat}}$ |
| $\alpha$ | 0.162 | 0.007 | 0.143 | 0.005 |
| $\beta$ | 3.126 | 0.073 | 3.218 | 0.071 |
| $\sigma_{\text{Ia}}$ | 0.102 | 0.006 | 0.086 | ⋯ |
| $\Delta_M$ | 0.088 | 0.013 | 0.044 | 0.010 |
| $\beta_0$ | 3.461 | 0.439 | 3.362 | 0.277 |
| $\beta_1$ | −1.542 | 1.440 | −0.660 | 1.145 |
| $\Delta_{M,0}$ | 0.065 | 0.019 | 0.054 | 0.020 |
| $\Delta_{M,1}$ | 0.082 | 0.054 | −0.038 | 0.060 |

**Note.** Nuisance parameters measured using the PSBEAMS method compared to those measured using the BBC method. $\beta_0$, $\beta_1$, $\Delta_{M,0}$, and $\Delta_{M,1}$ are intercept and slope parameters defining the linear $z$-dependence of $\beta$ and $\Delta_M$ (Equation (3)). $\beta$ evolution is detected with the PSBEAMS method but is an artifact of the analysis method as shown in Figure 12. Uncertainties on $\beta_0$ and $\beta_1$ are measured from the dispersion of simulations rather than the statistical uncertainties reported by the method.
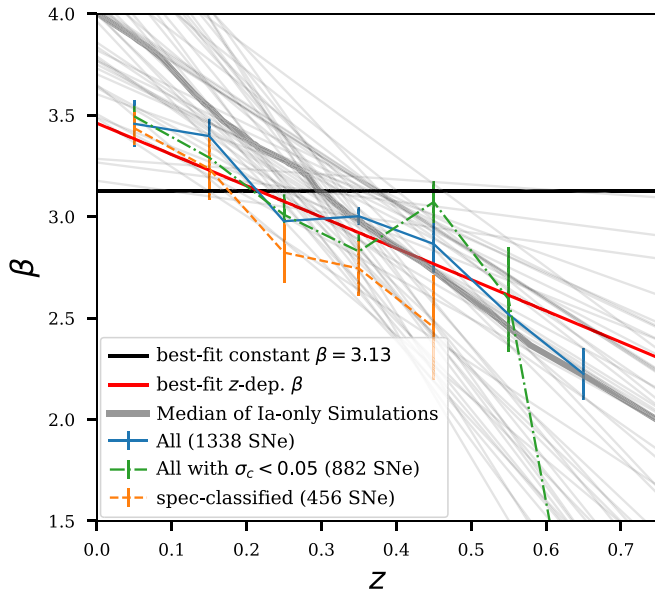
**Figure 12.** For the PSBEAMS method, evolution of the nuisance parameter $\beta$ in the full sample (blue), the full sample with a conservative $\sigma_C < 0.05$ (green), and the spectroscopically classified sample (orange). However, simulated SN samples (gray) with a constant simulated $\beta$ also show evidence for $\beta$ evolution. The BBC method does not find evidence for $\beta$ evolution, in either simulations or data.

Because the MDS sample has a median redshift of $\sim$0.35, the redshift range that can be used to constrain these parameters is limited. A large $\beta_1$ or $\Delta_{M,1}$ coefficient does not imply that a trend is observed to the maximum $z = 0.7$.

From MDS, CSP, and CfA SNe, J18 found a marginal (1.6$\sigma$) detection of $\beta$ evolution in the SN Ia data. In this analysis, using a more homogeneous data set, the PSBEAMS method gives a $\sim$3$\sigma$ detection of $\beta$ evolution, while the BBC method finds no evidence for $\beta$ evolution. For this reason, the trend is likely due to observational biases or an oversimplified analytic treatment. However, we investigated further for the PSBEAMS method by simulating SNe with a fixed input $\beta$ and then allowing $\beta$ to float in our likelihood model. The results of this test are shown in Figure 12, along with the measured $\beta$ evolution for both the full and spectroscopically classified samples.

The simulations have a wide dispersion in measured $\beta$ evolution, which may be due in part to the noisy high-$z$ data, or perhaps the limited redshift range of the sample. However, the general trend is negative, and our results from the data are consistent with these simulations, implying that the observed $\beta$ evolution is unlikely to be a physical effect. We therefore do not include it in our systematic uncertainty budget. Measurements of $\beta$ evolution with the BBC method do not have a statistically significant bias but do have wider dispersions than the measured 1$\sigma$ uncertainties by factors of 2−3 (Table 1).

We note that if $\beta$ is fixed as a function of redshift, the resulting value is not biased by the PSBEAMS analysis method (but can be biased by inaccurate priors when CC SN contamination is present). Similarly, the BBC framework explicitly corrects for both the bias in $\beta$ and the $z$-dependent bias in distance when $\beta$ is allowed to vary with redshift.

We do not measure significant evolution of the host galaxy mass step $\Delta_M$ with either BBC or the PSBEAMS method. However, we report $\beta(z)$ and $\Delta_M(z)$ in Table 1, and in Section 4.3 estimate what their contribution to the systematic uncertainty budget would have been had they been included in the error budget.

Hints of a nonzero $\beta(z)$ or $\Delta_M(z)$ were found in J18 and S18, respectively, while no evidence for an $\alpha(z)$ term has been found. There is also somewhat more physical motivation for $\beta$ and $\Delta_M$ evolution: dust properties are expected to evolve with redshift, and progenitor ages, which also evolve with redshift, may drive the $\Delta_M$ step (Childress et al. 2014). However, we also explore the possibility of $\alpha(z)$ here using the BBC method, finding that in our sample $\alpha(z)$ is only significant at the 1.5$\sigma$ level using both the G10 and C11 scatter models. We measure $\alpha(z) = 0.155 \pm 0.011 - 0.05740 \pm 0.038 \times z$ using the G10 model and a nearly identical step with the C11 model.

We note that if the SN Ia luminosity were evolving independent of changes in $\alpha$, $\beta$, or $\Delta_M$, we would be unable to distinguish this evolution from changes in $w$. However, if such evolution were physical, it would cause much larger discrepancies with $\Lambda$CDM, appearing to favor exotic dark energy, in studies with a larger redshift baseline such as DES Collaboration et al. (2019) or S18. No such effect has yet been observed.

### 4.2. Cosmological Parameters

In this section, we use SNe Ia in combination with external data sets to constrain three cosmological models:

1. A flat $\Lambda$CDM model.
2. The $w$CDM model. A redshift-independent $w$ is allowed to vary from the cosmological constant value of $w = -1$. We assume a flat universe.
3. The $w_a$CDM model. Using the Chevallier & Polarski (2001) and Linder (2003) formalism, $w$ is allowed to evolve with redshift: $w = w_0 + w_a z/(1 + z)$. We again assume a flat universe.

We use the more precise BBC distances to derive our baseline cosmological parameter measurements. BBC distance uncertainties are 18% smaller, on average, than PSBEAMS distance uncertainties (excluding the high-uncertainty $z = 0.01$ and $z = 0.7$ bins).

First, when assuming $\Lambda$CDM, we may derive useful, independent constraints on the cosmic matter density from SNe alone. We find the cosmic matter density $\Omega_m$ to be $0.353 \pm 0.037$ when assuming $\Lambda$CDM (the curvature, $\Omega_k$, is fixed at 0). This is consistent with the value of $\Omega_m = 0.315 \pm 0.007$ measured from the CMB (Planck Collaboration et al. 2018). J18 measured $0.319 \pm 0.040$ using the MDS high-$z$ data and the previous low-$z$ sample.

Next, we combine the binned BBC distances with the 2015 CMB likelihoods from *Planck* to constrain the $w$CDM model (Planck Collaboration et al. 2016; 2018 *Planck* likelihoods are currently unavailable). We find $w = -0.938 \pm 0.053$, consistent with a cosmological constant at the 1.2$\sigma$ level for a flat universe. Figure 13 shows the change in the $w$–$\Omega_m$ plane when using Foundation SNe as the low-$z$ sample instead of the previous low-$z$ SNe. From the spectroscopically classified data alone, we measure $w = -0.933 \pm 0.061$, which is consistent at the 1.1$\sigma$ level with $\Lambda$CDM and demonstrates that the use of photometrically classified data does not significantly change the measurement.

To increase the precision of these measurements, we include BAO constraints from Anderson et al. (2014), Ross et al. (2015), and Alam et al. (2017), which give measurements of

**Table 2**
$w$CDM and $w_a$CDM Parameters from MDS+Foundation SNe, BAO, CMB, and $H_0$

| | $w$CDM Constraints | |
|---|---|---|
| | $\Omega_m$ | $w$ |
| SNe+CMB | $0.331 \pm 0.017$ | $-0.938 \pm 0.053$ |
| SNe+CMB+BAO | $0.316 \pm 0.009$ | $-0.949 \pm 0.043$ |
| SNe+CMB+$H_0$ | $0.295 \pm 0.012$ | $-1.034 \pm 0.042$ |
| SNe+CMB+BAO+$H_0$ | $0.301 \pm 0.008$ | $-1.014 \pm 0.040$ |

| | $w_a$CDM Constraints | | |
|---|---|---|---|
| | $\Omega_m$ | $w_0$ | $w_a$ |
| SNe+CMB | $0.314 \pm 0.025$ | $-0.810 \pm 0.144$ | $-0.791 \pm 0.785$ |
| SNe+CMB+BAO | $0.321 \pm 0.010$ | $-0.825 \pm 0.095$ | $-0.570 \pm 0.401$ |
| SNe+CMB+$H_0$ | $0.280 \pm 0.011$ | $-0.734 \pm 0.082$ | $-1.541 \pm 0.374$ |
| SNe+CMB+BAO+$H_0$ | $0.305 \pm 0.008$ | $-0.895 \pm 0.095$ | $-0.597 \pm 0.439$ |

**Note.** Constraints on $w$CDM and $w_a$CDM using the BBC analysis method.
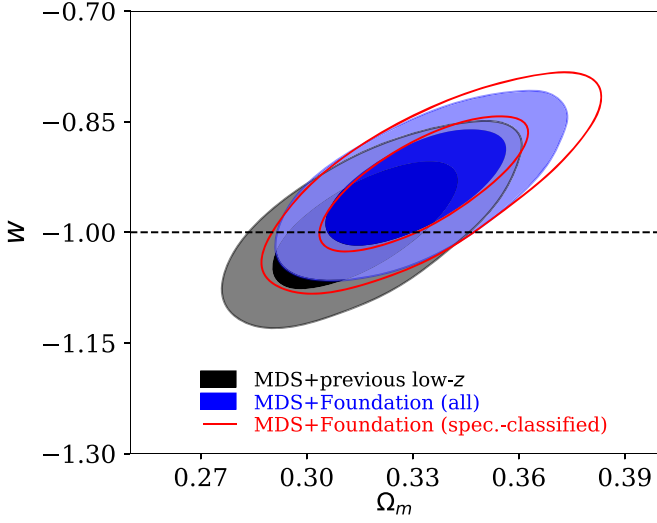


**Figure 13.** Constraints on $w$ and $\Omega_m$ assuming a flat $w$CDM model. The Foundation+MDS sample (blue) and the combined sample of Foundation and spectroscopically classified MDS SNe (red) are compared to the J18 results that use the previous low-$z$ data instead of Foundation (black).

**Table 3**
Summary of Systematic Uncertainties for $w$

| Uncertainty | MDS+Foundation | | MDS+Low-$z$ | |
|---|---|---|---|---|
| | | | (J18) | |
| | $\sigma_w^{\mathrm{sys}}$ | Rel. to $\sigma_w^{\mathrm{stat}}$ | $\sigma_w^{\mathrm{sys}}$ | Rel. to $\sigma_w^{\mathrm{stat}}$ |
| All sys. | 0.041 | 1.241 | 0.043 | 1.144 |
| Phot. cal. | 0.027 | 0.832 | 0.019 | 0.495 |
| $-$ SALT2 model | 0.023 | 0.699 | 0.008 | 0.203 |
| $-$ PS1 cal. | 0.016 | 0.496 | 0.007 | 0.190 |
| CC contam. | 0.013 | 0.381 | 0.013 | 0.334 |
| Bias corr. | 0.011 | 0.340 | 0.020 | 0.520 |
| MW $E(B-V)$ | 0.007 | 0.205 | 0.014 | 0.379 |
| Pec. vel. | 0.006 | 0.181 | 0.007 | 0.197 |
| Mass step | 0.000 | 0.000 | 0.018 | 0.469 |

**Note.** Each systematic uncertainty as a fraction of the statistical uncertainties for the MDS+low-$z$ sample ($\sigma_{\mathrm{stat}} = 0.038$) and the MDS+Foundation sample ($\sigma_{\mathrm{stat}} = 0.034$).
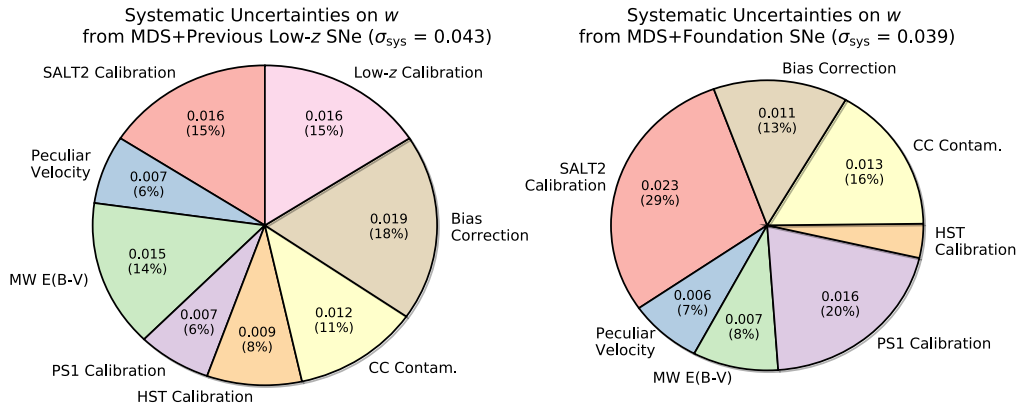
best measurement of SNe+CMB+BAO+$H_0$ is within $1\sigma$ of $\Lambda$CDM.

### 4.3. Systematic Uncertainties

The systematic uncertainties from this analysis compared to the J18 analysis are shown in Table 3 and Figure 14 using the BBC framework. A full table of systematic uncertainties for both the full and spectroscopically classified samples, using both PSBEAMS and BBC, is shown in Table 5. We focus on the systematic uncertainties from the BBC method here and discuss the difference between the methods in Section 4.4.

With the reduction in the distance bias systematic uncertainty and the use of just a single sample, photometric calibration is the dominant systematic uncertainty (0.027) in this sample. This systematic uncertainty can be split into three components: the SALT2 calibration, *HST* CALSPEC calibration, and uncertainty in the PS1 calibration. The SALT2 and PS1 calibrations are the dominant components of the systematic uncertainty budget.

The bias correction systematic uncertainty has been reduced from 0.02 in J18 to just 0.011 in the current analysis, a result of the better-understood selection effects in Foundation. The

the BAO scale at $z = 0.15, 0.32, 0.38, 0.51, 0.57$, and 0.61. The BAO scale is proportional to a combination of the angular diameter distance to a given redshift and the Hubble parameter at that redshift and is measured from the signature of acoustic waves on the cosmic matter distribution. We also use $H_0$ constraints from Riess et al. (2018), which shift the measured value of $w$ to be $\sim$6% more negative than SN+CMB+BAO owing to their $3.7\sigma$ inconsistency with the *Planck* results.

The full constraints on the $w$CDM model from this combination of different probes are given in Table 2. For the $w$CDM model, including BAO constraints moves $w$ slightly closer to $\Lambda$CDM, but still $1.2\sigma$ from $w = -1$, and including the local $H_0$ constraints moves $w$ to $-1.014 \pm 0.040$ (however, the CMB and local $H_0$ measurements are internally inconsistent).

Allowing $w$ to evolve with redshift gives $w_0 = -0.810 \pm 0.144$ and $w_a = -0.791 \pm 0.785$ from SNe+CMB, which is consistent with $\Lambda$CDM. Including $H_0$ constraints moves the value to nearly $3\sigma$ from $\Lambda$CDM, due to the internal inconsistency of local $H_0$ measurements with *Planck*, but the

**Figure 14.** Systematic uncertainties on $w$ from J18 (left) compared to this analysis (right). The size of each chart is proportional to the size of the total systematic uncertainty budget for each analysis. The size of each slice corresponds to the size of each systematic uncertainty as a fraction of the sum of all systematic uncertainties. The calibration systematic uncertainties are 42% higher in this analysis, as we have just one photometric system, but the bias correction systematic uncertainty is 55% lower. The overall systematic uncertainty is 9% lower.

**Table 4**
Measurements of $w$ from Alternative Methods of Marginalizing over CC SNe

|  | $w$ | $\Delta w$ |
|---|---|---|
| Baseline | $-0.920 \pm 0.033$ | ... |
| CC SN simulations | $-0.924 \pm 0.033$ | $-0.004$ |
| CC SN prior | $-0.938 \pm 0.033$ | $-0.018$ |
| Classification prior[a] | $-0.886 \pm 0.036$ | $0.034$ |
| Nuisance parameters fixed | $-0.922 \pm 0.033$ | $-0.002$ |

**Notes.** For the BBC method, changes in $w$ (statistical uncertainties alone) after applying analysis variants related to the use of CC SN-contaminated data (see Table 7 of J18 for a similar table for the PSBEAMS method). We show the effect of using CC SN simulations with alternate LFs, dust distributions, or CC SN templates, the effect of using the parametric (Hlozek et al. 2012) prior on the CC SN distribution, the effect of a different light-curve classifier, and the effect of fixing the nuisance parameters to the values derived from the spectroscopically classified data.
[a] The large change in $w$ is due to the highest two redshift bins, which have CC SNe significantly blended with SNe Ia and are de-weighted when all systematic uncertainties are applied. The PSBEAMS method generally performs more consistently than BBC with the PSNID classifier, as it has an additional parameter to scale the "overconfident" $P(\text{Ia})$ probabilities (Jones et al. 2017).

uncertainty due to marginalizing over CC SN contamination remains approximately the same at ~1.3%. This systematic uncertainty remains subdominant to photometric calibration and will be improved in future work. Statistical-uncertainty-only measurements of $w$ after applying a number of different treatments of the CC SN contamination are shown in Table 4. These include alternate light-curve classification methods, predicted biases from simulations, alternate parameterizations of the contaminating distribution, and nuisance parameters forced to be equal to those measured from the spectroscopically classified data.

Although we do not include possible redshift dependence of the mass step or $\beta$ in our systematic uncertainty budget, we show their effect in Table 5. Compared to the PSBEAMS method, BBC bias corrections in shape and color bring the high-$z$ mass step, $\Delta_M = 0.043 \pm 0.013$, in better agreement with the low-$z$ mass step, $\Delta_M = 0.060 \pm 0.024$, lowering the measured systematic uncertainty due to an evolving mass step.

We find that $\beta$ evolution is also a slightly lower systematic uncertainty in the BBC method for similar reasons.

### 4.4. Comparing BBC to the PSBEAMS Method

Use of the BBC method gives a more fine-grained approach to bias corrections, as parameters $m_B$, $x_1$, and $c$ are each corrected for selection biases. The results are consistent with the values from the PSBEAMS method, but the final uncertainty on $w$ is reduced by 7%. We find that the final value of $w$ is lower by 0.020 when using the BBC method; however, the value when excluding systematic uncertainties is 0.041 *higher* when using the BBC method in the highest-redshift bins. The BBC results on spectroscopically classified data alone are somewhat closer to results using the PSBEAMS method; $w$ is higher by 0.022 when using BBC compared to the $w$ from the PSBEAMS method.

For the full sample, the difference in statistics-only $w$ values, as well as the relatively large change of 0.043 between the PSBEAMS measurement with statistical uncertainties only and the PSBEAMS measurement when systematic uncertainties are included, is primarily due to systematic uncertainties in marginalizing over CC SNe. Use of the NN classifier compared to the PSNID classifier to assign the prior probabilities that a given SN is of Type Ia shifts $w$ by $+0.028$ (although, due to significant covariances between bins, the uncertainty on $w$ does not increase by the same amount).

With BBC, the method of marginalizing over CC SNe is somewhat different. First, we use a different nominal classification method (Appendix A.2). The PSBEAMS method also includes a parameter that allows probabilities to be shifted linearly (see Equation (5)), while BBC does not. Finally, the PSBEAMS method uses a point-to-point, parameterized linear interpolation of the CC SN distribution, while BBC uses the simulations themselves as a fixed prior on the distance. The two methods yield distances consistent to $\ll 1\sigma$ at $z < 0.4$ but disagree somewhat at high $z$ as shown in Figure 15. CC SNe in the $z \gtrsim 0.5$ redshift range can be difficult to marginalize over, as the CC SN distribution begins to overlap substantially with the SN Ia distribution because of Malmquist bias. As these methods were developed and tested independently, they are complimentary methods for measuring $w$ from photometrically classified data.

Table 5
Summary of $w$ Measurements and Systematic Uncertainties

| | Full Sample | | | | Spec Sample | | | | J18 | |
|---|---|---|---|---|---|---|---|---|---|---|
| | PSBEAMS | | BBC | | PSBEAMS | | BBC | | | |
| $w_{\mathrm{stat+sys}}$ | $-0.918 \pm 0.057$ | | $-0.938 \pm 0.053$ | | $-0.955 \pm 0.063$ | | $-0.933 \pm 0.061$ | | $-0.990 \pm 0.057$ | |
| $w_{\mathrm{stat}}$ | $-0.961 \pm 0.036$ | | $-0.920 \pm 0.033$ | | $-0.954 \pm 0.050$ | | $-0.936 \pm 0.047$ | | $-1.022 \pm 0.038$ | |
| Sys. Error | $\Delta w$ | $\Delta \sigma_w$ | $\Delta w$ | $\Delta \sigma_w$ | $\Delta w$ | $\Delta \sigma_w$ | $\Delta w$ | $\Delta \sigma_w$ | $\Delta w$ | $\Delta \sigma_w$ |
| Photometric calibration | 0.026 | 0.037 | −0.020 | 0.027 | 0.013 | 0.037 | 0.009 | 0.026 | 0.012 | 0.019 |
| − SALT2 model[a] | 0.029 | 0.026 | −0.015 | 0.023 | 0.006 | 0.026 | −0.000 | 0.021 | 0.000 | 0.008 |
| − Supercal | 0.001 | 0.001 | −0.003 | 0.000 | −0.001 | 0.000 | 0.000 | 0.000 | −0.001 | 0.000 |
| − Filter functions | −0.001 | 0.007 | −0.005 | 0.000 | 0.004 | 0.008 | 0.002 | 0.000 | −0.001 | 0.009 |
| − PS1 zero-point | −0.005 | 0.024 | −0.011 | 0.016 | 0.003 | 0.026 | 0.008 | 0.017 | −0.002 | 0.007 |
| − *HST* calibration | 0.001 | 0.008 | −0.002 | 0.003 | 0.003 | 0.000 | 0.002 | 0.005 | 0.002 | 0.009 |
| Mass step | 0.012 | 0.009 | −0.001 | 0.000 | −0.000 | 0.000 | 0.001 | 0.000 | 0.005 | 0.018 |
| CC contamination | 0.040 | 0.017 | −0.000 | 0.013 | 0.000 | 0.000 | 0.000 | 0.000 | −0.001 | 0.013 |
| Bias correction | 0.012 | 0.009 | −0.009 | 0.011 | −0.007 | 0.004 | −0.008 | 0.019 | 0.012 | 0.020 |
| Peculiar velocity | −0.002 | 0.010 | −0.002 | 0.006 | 0.002 | 0.007 | 0.004 | 0.007 | 0.002 | 0.007 |
| MW $E(B-V)$ | −0.000 | 0.008 | −0.002 | 0.007 | 0.000 | 0.000 | 0.003 | 0.010 | 0.009 | 0.014 |
| $\beta(z)$[b] | −0.008 | 0.025 | 0.004 | 0.018 | 0.012 | 0.036 | −0.004 | 0.012 | 0.012 | 0.016 |
| $\Delta_M(z)$[b] | 0.039 | 0.027 | −0.001 | 0.000 | −0.000 | 0.000 | 0.003 | 0.000 | 0.006 | 0.020 |
| $\beta(z)$[b] | −0.008 | 0.024 | 0.004 | 0.018 | 0.012 | 0.036 | −0.004 | 0.012 | 0.012 | 0.016 |
| $\Delta_M(z)$[b] | 0.040 | 0.026 | −0.001 | 0.000 | −0.000 | 0.000 | 0.003 | 0.000 | 0.006 | 0.020 |

**Notes.** Summary of measurements and systematic uncertainties for each value of $w$ presented in this paper.
[a] J18 reported a SALT2 systematic uncertainty that was smaller than the one used in this analysis and the Pantheon analysis. Corrected, a SALT2 systematic shift with the same size as the one used in this analysis gives a systematic uncertainty of 0.016 for the J18 analysis instead of the reported 0.008. This increase does not significantly impact the final constraints on $w$ from J18.
[b] As the $z$-dependence of $\beta$ and mass step have not been significantly detected in any analysis to date, we have not included them in the final systematic uncertainty error budget (also following J18 and the Dark Energy Survey analysis; Brout et al. 2019). However, we show their impact here. Including them would not significantly increase the final BBC measurement uncertainty but would increase the uncertainties on $w$ from the PSBEAMS method by 25%.
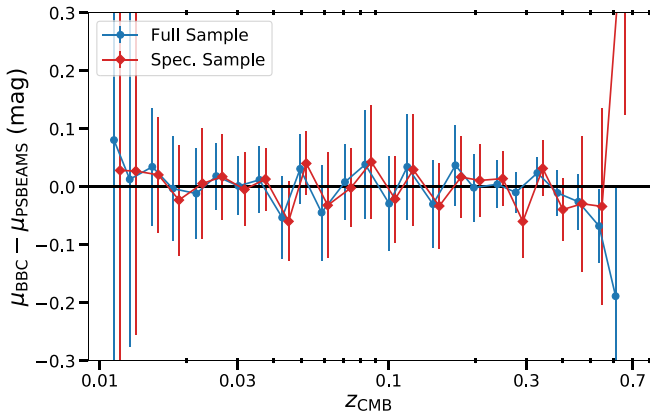


**Figure 15.** Difference between distances derived using the PSBEAMS and BBC methods, for both the full sample (blue) and spectroscopically classified sample (orange), as a function of redshift. Some modest discrepancies occur at high $z$, where the only detectable CC SNe have brightnesses that are close to or greater than those of SNe Ia. Points have been offset slightly for visual clarity.

## 5. Discussion and Future Directions

### 5.1. Differences between Foundation and Other Low-z Samples

The measurement of $w$ presented here is independent of all previous low-$z$ data sets and most previous high-$z$ data sets. Figure 16 compares the measurement of $w$ after including the Foundation sample to other measurements of $w$ published within the past 5 yr. All previous measurements use some combination of CfA and CSP data at low $z$, and the Betoule et al. (2014) and S18 analyses also use the same SDSS, SNLS (Astier et al. 2006), and *HST* data. J18 used the same high-$z$

sample as this analysis does, and after replacing the previous low-$z$ compilation with the Foundation sample, we find that $w$ is shifted by +5%.

The total systematic uncertainty of our measurement of $w$ is 9% smaller than in the J18 analysis; although the total calibration systematic uncertainty is increased by 42%, the selection bias systematic uncertainties are reduced by 55% (Figure 9). This calibration systematic uncertainty is due in approximately equal parts to the calibration of the SALT2 model and the PS1 zero-point calibration uncertainties. The uncertainties due to survey calibration in particular are larger because the previous low-$z$ sample was observed using $\geqslant 3$ filters per telescope on up to seven different photometric systems, while this analysis uses only two filters on a single photometric system.

Reducing these systematic uncertainties will require adding SNLS, SDSS, or DES high-$z$ data to the analysis; adding additional low-$z$ samples, though these may have lower weight than the Foundation sample when systematic uncertainties are included; or using a retrained SALT2 model. Because SN surveys are not consistently calibrated before training the SALT2 model, a retrained SALT2 model that includes Supercal-corrected photometry (Scolnic et al. 2015) would substantially improve future analyses. Similarly, a SALT2 model trained to use redder rest-frame wavelengths where SNe Ia are better standardizable candles (e.g., Mandel et al. 2011) would also improve distance measurements and improve the ability to plan future near-IR SN surveys such as *WFIRST* (Pierel et al. 2018). Though the MDS data are redshifted enough for all four filters to be used, 50% of Foundation data (the $iz$ observations) are not used in this work. For this reason
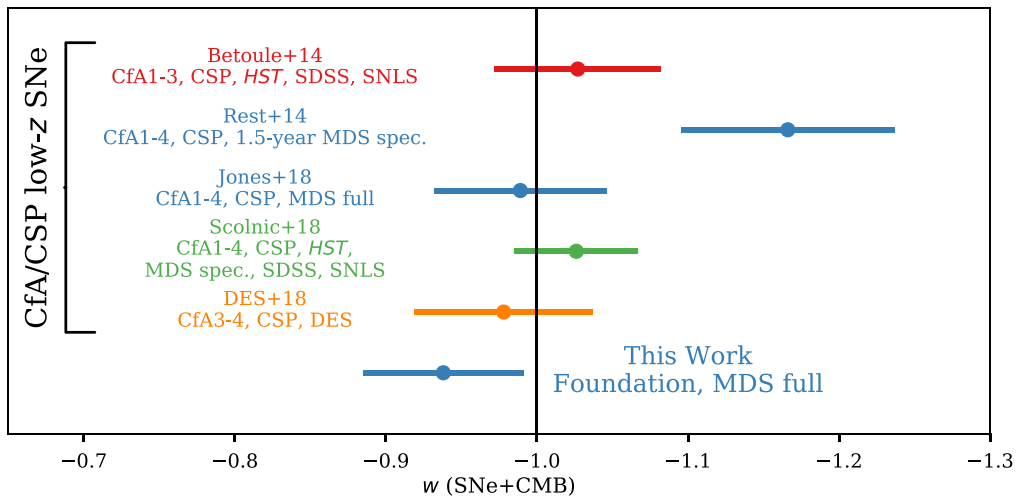
**Figure 16.** Measurements of $w$ using SNe with CMB constraints from *Planck*, published within the past 5 yr. From top to bottom, we show measurements from JLA (Betoule et al. 2014), spectroscopically classified SNe Ia from the first 1.5 yr of the MDS (Rest et al. 2014), the full MDS sample (J18), the Pantheon sample (S18), spectroscopically classified SNe Ia from the first 3 yr of DES (DES Collaboration et al. 2019), and the present analysis. The previous analyses share many of the same low-$z$ and high-$z$ SN samples, which we indicate in the figure. The "spec." abbreviation indicates that only spectroscopically classified MDS SNe were used in these analyses.

alone, an SN light-curve fitter with redder wavelength coverage would provide enormous benefits by (1) reducing the sample dispersion by using data at wavelengths where SNe Ia are better standard candles and (2) reducing the single-filter calibration systematic uncertainties by including twice as many filters at low $z$.

The Foundation sample is much more similar to the high-$z$ sample—in sample selection, photometric reduction, and photometric system—than the previous low-$z$ samples. With this sample, we therefore expect any bias caused by unforeseen systematic uncertainties (e.g., unexpected dependence of SNe Ia on their host galaxies) to be greatly reduced. Unforeseen systematic uncertainties would increase the error on our measurement but would affect previous measurements more strongly. Although the shift in distances with these new data is only marginally significant, it may well be that this shift is hinting at an aspect of SN Ia physics that will be revealed in future work. The slight change in $w$ found in this work is driven by the low-$z$ sample, which also has SNe with masses and sSFRs that are significantly shifted with respect to the previous low-$z$ sample (Figure 3).

The Foundation Supernova Survey aims to observe up to 800 cosmologically useful SNe Ia. With such large statistical leverage, we may be able to better understand the ways in which SN Ia distances may be affected by unexpected systematic uncertainties.

### 5.2. Photometric Classification

We have shown that the $\sim 1\sigma$ difference between the value of $w$ measured from this data and the previous (J18) results stems from the low-$z$ sample alone and not from any biases caused by marginalizing over CC SNe. First, our results are consistent at $\ll 1\sigma$ with the spectroscopically classified data. Second, we have employed five separate approaches to modeling the CC SN distribution: the PSBEAMS method uses a single-Gaussian CC SN model, a two-Gaussian CC SN model, and an asymmetric Gaussian model, each of which has means and dispersions that are linearly interpolated between five control points across the redshift range of the sample. With the BBC

method, we use a Monte Carlo simulation-based prior and a parameterized single-Gaussian CC SN model that evolves as a second-order polynomial across the redshift range. Finally, we adopt three different methods of classifying CC SNe. All methods yield results $\sim 1\sigma$ from $\Lambda$CDM.

CC SN contamination is not the dominant systematic uncertainty in this work, but it remains a significant one at 1.3%. It will need to be reduced in future work, and the true distribution of CC SNe will need to be better understood. However, we see no evidence, either in this work or from previous recent studies of using photometrically classified SNe for cosmology, that our conservative estimation of this systematic uncertainty is unrealistic.

An additional consideration related to CC SN marginalization was found by Knights et al. (2013), who note that the BEAMS formalism breaks down in samples with large correlated systematic uncertainties. However, their analysis explored correlations on the order of $\sim 10\%$, and only percent-level correlations exist in this sample. We expect that a more sophisticated treatment of systematic uncertainties will only become necessary in analyses with even larger samples (e.g., LSST).

### 5.3. The Relationship between SNe Ia and Their Host Galaxies

The relationship between SNe Ia and their host galaxies is subject to significant uncertainty and could bias cosmological parameters (e.g., Childress et al. 2014; Rigault et al. 2018). To mitigate this uncertainty, we explored several methods of estimating the potential bias to $w$ from the uncertain relation between SNe Ia and their host galaxies. First, in Section 4 we examined the effect of allowing the host galaxy mass step to evolve with redshift, as predicted by Childress et al. (2014) and observed with marginal significance by S18 (although not detected by J18). In this work we have a limited redshift range over which the mass step can evolve, and although we do not find significant evidence for mass step evolution, the uncertainties are much higher than in the Pantheon analysis of S18. Allowing an evolving mass step does not shift the measurement of $w$ when the BBC analysis method is used
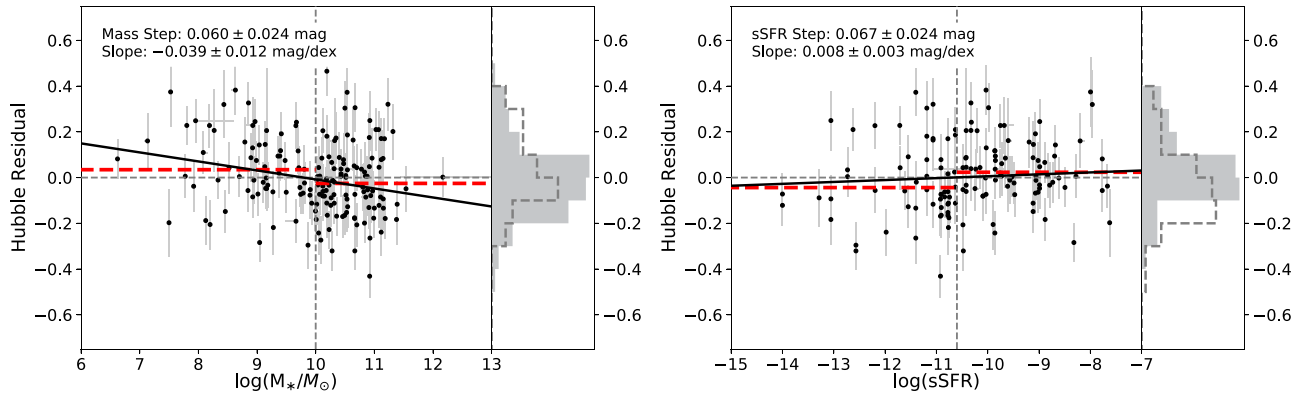
**Figure 17.** For the Foundation sample, the correlation of Hubble residuals with host galaxy mass (left) and Hubble residuals with host galaxy sSFR (right). Assuming a step function to describe the sample, red dashed lines show the maximum likelihood average Hubble residual for each side of the step, which is indicated by a vertical gray dashed line. The solid black line represents the best-fitting linear function. In the left panel, histograms show the Hubble residuals of SNe in low-mass (dashed line) and high-mass (filled) host galaxies. In the right panel, histograms show the Hubble residuals of SNe in low-sSFR (dashed) and high-sSFR (filled) host galaxies. We use the BBC method to generate the Hubble residuals shown here.

(Table 5). The PSBEAMS measurement is shifted owing to observational biases as discussed in Section 4.1. Given that the spectroscopically classified data do not prefer an evolving mass step, we expect that the large $\Delta_M$ systematic uncertainty when using the PSBEAMS method is not a physical effect.

Second, a number of recent papers have suggested alternative relationships between SN Ia distance measurements and their host galaxy masses. For example, some evidence has also shown that metallicity, a function of both mass and SFR, may correlate with Hubble residuals better than mass alone (Hayden et al. 2013), a possible systematic uncertainty that warrants investigation with additional data. For the Foundation sample, the correlation of Hubble residuals with host galaxy mass and sSFR is shown in Figure 17.

Recent work has also explored the relationship between SN Ia Hubble residuals and host galaxy properties near the SN location. Within the past year, Rigault et al. (2018) found a strong relationship between the SN distance measurement and the sSFR near the SN location using SNFactory data, Roman et al. (2018) found a strong relationship between the SN distance and the local (and global) rest-frame $U - V$ color of the host, and Jones et al. (2018a) found a $3\sigma$ relationship between the SN distance and the local host galaxy mass *after* global mass correction. In the Jones et al. (2018a) data, the relationship between the SN distance and local host galaxy mass was larger than either the sSFR step or a local color step, and the local mass step was 60% larger than the alternate steps after global mass correction. As our knowledge of the relationship between SN Ia distances and host galaxies becomes increasingly fine-grained and complex, it is important to test for these effects in cosmological analyses.

Although we cannot measure host galaxy properties within ~1–2 kpc of the SN location for the $z > 0.1$ SN sample owing to seeing limitations, we estimate the bias due to a possible local mass or sSFR step in two ways. First, Jones et al. (2018a) found that a step using the sSFR measured from the global properties of the host galaxy was slightly more significant than a local sSFR step, and so we measure $w$ by replacing the global mass step with a global sSFR step. The sSFR values are estimated using the method discussed in Section 2.2, but we reduce the sample slightly by using only SNe Ia with SDSS $u$-band data to ensure robust SFRs. For the higher-$z$ SNe, this restriction is unimportant, but the majority of our sample is not

at high enough redshift to provide rest-frame $u$. Including SNe for which robust sSFR measurements cannot be computed, this restriction reduces our sample by ~300 SNe. We then use both the BBC and PSBEAMS likelihoods to estimate the distances after using an sSFR step instead of a mass step.

For the BBC method, we find an insignificant sSFR step of $\Delta_{sSFR} = 0.018 \pm 0.011$ mag for the G10 scatter model and a similarly small step for the C11 model. Fitting with an undetected sSFR step instead of the detected mass step therefore shifts $w$ in a way that is not justified by the data (a shift of $+0.049$). For the PSBEAMS method, we find a larger sSFR step of $\Delta_{sSFR} = 0.038 \pm 0.013$, but fitting for an sSFR step instead of a mass step gives a statistics-only measurement of $w$ that is shifted by just $-0.014$. Jones et al. (2018a) found a somewhat larger sSFR step from low-$z$ data alone of $0.054 \pm 0.020$; it may be that the higher-$z$ measurements with more limited wavelength coverage have less accurate sSFR measurements, but the values of these two steps are statistically consistent. Both are also consistent with the sSFR step found by Brout et al. (2019) of $0.037 \pm 0.025$ mag. We therefore do not expect that use of the mass step instead of an sSFR step is biasing our measurements of $w$.

Second, we can compute $w$ using only SNe in locally massive regions of their host galaxies. From the Jones et al. (2018a) public data, 83% of SNe in globally massive host galaxies that occur at host $R < 2$ (near their host center; $R$ is discussed in Section 2.2) are in locally massive regions. This subset cuts the sample size by ~40%, but if a strong change in $w$ is observed with this subset, it may mean that measurements of dark energy are sensitive to the relationship between SNe Ia and their local host galaxy environments. Using the BBC method, $w$ is shifted by $-0.029$, a shift equal to ~90% of the statistical uncertainty for the full sample. For the PSBEAMS method, however, $w$ is changed by less than 0.001.

These three tests are summarized in Table 6. They show that $w$ is not sensitive to a mischaracterization of the relationship between SNe Ia and their host galaxies, although the statistically insignificant, ~3% change in $w$ when controlling for a local mass step may warrant investigation in future work.

## 6. Conclusions

We combine 1164 SNe Ia from the Pan-STARRS1 medium deep survey with 174 SNe Ia from the Foundation Supernova

**Table 6**
Alternate Relationships between SN Ia Hubble Residuals and Host Galaxy Properties and Their Effect on $w$

| | PSBEAMS | | BBC | |
|---|---|---|---|---|
| | Step Size | $\Delta w$ | Step Size | $\Delta w$ |
| Default mass step | $0.088 \pm 0.013\ (6.7\sigma)$ | ⋯ | $0.044 \pm 0.010\ (4.4\sigma)$ | ⋯ |
| $z$-dependent mass step | $\Delta_{M,0} = 0.065 \pm 0.019\ (3.4\sigma)$ $\Delta_{M,1} = 0.082 \pm 0.054\ (1.5\sigma)$ | 0.032 | $\Delta_{M,0} = 0.054 \pm 0.020\ (2.7\sigma)$ $\Delta_{M,1} = -0.038 \pm 0.060\ (0.6\sigma)$ | 0.016 |
| Global sSFR step[a] | $0.038 \pm 0.013\ (2.9\sigma)$ | $-0.014$ | $0.018 \pm 0.011\ (1.6\sigma)$ | 0.049 |
| SNe in locally massive regions[b] | ⋯ | $-0.001$ | ⋯ | 0.029 |

**Notes.**
[a] The 0.049 shift in $w$ when using the BBC method is not physical, as we replace a highly significant mass step correction with a marginally significant sSFR step.
[b] PS1 does not have sufficient resolution to measure the local mass step at high $z$, so we restrict to SNe in probable locally massive regions. Jones et al. (2018a) measured the local mass step to be $0.067 \pm 0.017$ mag.

Survey DR1 to measure cosmological parameters from a single telescope and photometric system. The well-calibrated PS1 system makes this an excellent unified sample for precision cosmology. Future Foundation data releases will build the SN Ia sample size up to 800 SNe Ia at $z < 0.1$.

Similar to the high-$z$ sample, Foundation is nearly an untargeted, magnitude-limited sample, similar to the higher-$z$ data. Foley et al. (2018) note that 86% of the DR1 sample was independently discovered by the untargeted ASAS-SN and PSST surveys (though the untargeted surveys were not the first surveys to discover 25% of those SNe). Including discoveries from other untargeted surveys brings the total to 94%. In sample selection, photometric system, and host galaxy demographics (Figure 3), Foundation is therefore a more homogeneous sample across the redshift range than previous data sets. One caveat is that the photometrically classified MDS sample is biased toward more luminous host galaxies owing to the necessity of obtaining a spectroscopic redshift. A future analysis with photometric redshifts could ameliorate this concern.

Foundation distances are $0.046 \pm 0.027$ mag fainter than distances from the previous low-$z$ data (1.7$\sigma$ significance). After combining our SN Ia distances with *Planck* priors, we find $w = -0.938 \pm 0.053$, an uncertainty 7% lower than the previous J18 analysis that used only the CfA and CSP low-$z$ compilations. Allowing $w$ to evolve with redshift using the parametric form given by Linder (2003) gives $w_0 = -0.810 \pm 0.144$ and $w_a = -0.791 \pm 0.785$.

The systematic uncertainty attributed to SN selection biases has decreased by 55%. This decrease is due to the well-understood selection criteria of the Foundation Supernova Survey and the similarity of these criteria to previous high-$z$ analyses. The photometric calibration uncertainty has increased relative to the previous analysis by 42% as fewer independent photometric systems are used, but adding additional high-$z$ or low-$z$ data to this analysis, or using a SALT2 model retrained with redder rest-frame data, will reduce this source of uncertainty.

All Foundation SNe also have observations in $i$ and $z$ filters, wavelengths at which SNe Ia are better standard candles. Currently, the SALT2 model extends redward only to 7000 Å, and therefore SALT2 cannot fit these data. Given that the Foundation data extend to rest-frame $z$ band, they could be used to retrain the SALT2 model on redder data and maximize the utility of this data set.

In many ways the Foundation data allow us to remove substantial uncertainty regarding SN Ia physics owing to the similarity of the PS1-observed SN data across the redshift range. The 1.7$\sigma$ discrepancy between Foundation data and previous low-$z$ data may be a statistical fluctuation but could also indicate that unforeseen systematic uncertainties, related to either photometric calibration or SN physics, are affecting the data.

Many additional SN Ia survey telescopes are currently collecting data or will be in the near future (DES, DES Collaboration et al. 2019; LSST, The LSST Dark Energy Science Collaboration et al. 2018; *WFIRST*, Hounsell et al. 2018). Most or all of these missions will have large samples of photometrically classified SNe, and improving some of the methods discussed here to marginalize over CC SN contamination will be vital to the accuracy of future cosmological constraints. However, these surveys will not observe large samples of low-$z$ SNe. With every subsequent addition to the Foundation sample, the Hubble diagram anchor will improve in statistical precision and the properties of dark energy will be better understood.

## Appendix
## The BEAMS Likelihood Model

In its simplest formulation, BEAMS models SN Ia and CC SN distances with Gaussian likelihoods where the means of the Gaussians—the distance, or average distance within a redshift bin—are free parameters. From Kunz et al. (2007) and using the notation in Kessler & Scolnic (2017) and J18, the probability of the model given the data, $D$, is

$$P(\theta|D) \propto P(\theta) \times \prod_{i=1}^{N} (\mathcal{L}_i^{\text{Ia}} + \mathcal{L}_i^{\text{CC}}),$$

$$\mathcal{L}_i^{\text{Ia}} = \tilde{P}_i(\text{Ia}) D_{\text{Ia}}(z_i, \mu_i, \mu_{\text{model},i}),$$

$$D_{\text{Ia}} = \frac{1}{\sigma_\mu \sqrt{2\pi}} \exp[-\chi_{\text{HD}}^2/2]. \quad (4)$$

$(P\theta|D)$ is the posterior probability of the model given the data, $D$, which is proportional to the priors on free parameters $\theta$ and the product of the likelihoods for the $N$ SNe in the sample. For the $i$th SN, $\mathcal{L}_i^{\text{Ia}}$ and $\mathcal{L}_i^{\text{CC}}$ are the likelihoods of the SN Ia and CC SN distribution, respectively. The CC SN likelihood, $\mathcal{L}_i^{\text{CC}}$, is identical to the form of the Type Ia likelihood in Equation (4), except that the $\mu_{\text{model},i}$ term is now the related to the brightness of CC SNe, shape- and color-corrected as though they were SNe Ia, rather than the cosmological distance modulus derived from SN Ia standardization. $\tilde{P}_i(\text{Ia})$ is the adjusted prior probability that the $i$th SN is of Type Ia (see Equation (5) below), and $\chi_{\text{HD}}^2$ is the $\chi^2$ of the model distance compared to the measured distance.

In the methods outlined below, $\mu_{\text{model},i}$ is allowed to depend on the redshift, and distances are computed by including the Tripp formula, Equation (2), in the likelihood. The Tripp formula, in turn, depends on the global free parameters $\alpha$, $\beta$, and $\Delta_M$. The PSBEAMS and BBC implementations of BEAMS are discussed in much greater detail in J18 and Kessler & Scolnic (2017), respectively. However, we give a broad outline of both methods, as well as the differences between them, below.

### A.1. The PSBEAMS Method

In the PSBEAMS likelihood, $\mu_{\text{model}}$ is linearly interpolated between a series of $\log(z)$-spaced control points $z_b$ across the redshift range (0.01, 0.7). Betoule et al. (2014) find that the difference between this approximation and $\Lambda$CDM is always smaller than 1 mmag across the $0.01 < z < 1.3$ redshift range (see also Marriner et al. 2011, who first used redshift-binned distances that were independent of cosmological parameters). PSBEAMS allows the dispersion of SNe Ia, $\Sigma_{\text{Ia}}$, to be a free parameter, such that in Equation (4) we have $\sigma_\mu^2 = \sigma_i^2 + \Sigma_{\text{Ia}}^2$. The standard deviation of the SN Ia likelihood—approximately equivalent to the intrinsic dispersion of the SN Ia population—is held constant as a function of redshift. The CC SN

population, however, has a mean and dispersion that evolve with redshift according to the same linear interpolation method as $\mu_{\text{model}}$. The SN Ia and CC SN components of the likelihood are multiplied by prior probabilities (discussed in the previous section) that a given SN is a CC SN or a Type Ia. As this is an abbreviated discussion of the PSBEAMS likelihood model, we direct the reader to the full mathematical formalism, which is given in J18, Section 3.1.

The relation between SN Ia distance and host galaxy mass is modeled by using two Gaussian distributions to model the SN Ia likelihood, one for SNe Ia in high-mass host galaxies and one for those in low-mass host galaxies (with a default high/low-mass boundary at $\log(M_*/M_\odot) = 10$). $\mathcal{L}_i^{\text{Ia}}$ then becomes $\mathcal{L}_i^{\text{Ia},M<10} + \mathcal{L}_i^{\text{Ia},M>10}$. Each of these two likelihoods includes the prior probability that a given host galaxy is high- or low-mass, and those probabilities are given by the (assumed Gaussian) uncertainties on host mass from the method discussed in Section 2.2.

In total, the baseline likelihood model has 41 free parameters. These parameters include the mean of the SN Ia distribution at 25 control points, $\mu_{\text{model}}(z_b)$, which is equivalent to the SN Ia distance measurement. The parameters also include the mean and standard deviation of the CC SN distribution at five control points (10 parameters). Additionally, the Tripp parameters $\alpha$ and $\beta$, the host galaxy mass step $\Delta_M$, and the standard deviation of the SN Ia distribution $\Sigma_{\text{Ia}}$ (not a function of redshift) are all free parameters. Finally, we allow the prior probabilities that an SN is Type Ia to be linearly shifted and renormalized (two parameters). The relationship between the normalization factor, $A$, the shift parameter, $S$, and the adjusted probabilities for the $i$th SN is given by

$$\tilde{P}_i(\text{Ia}) = \frac{A \times (P_i(\text{Ia}) + S)}{1 - (P_i(\text{Ia}) + S) + A \times (P_i(\text{Ia}) + S)}$$
$$0 < \tilde{P}_i(\text{Ia}) < 1. \quad (5)$$

We apply loose priors on all parameters except the mean of the SN Ia distribution $\mu_{\text{model}}(z_b)$, so as not to impose priors on the SN Ia distances, and sample the posterior with a Markov chain Monte Carlo algorithm (the `emcee` Parallel-tempered Ensemble sampler; Foreman-Mackey et al. 2013). Additional details about this procedure are given in Jones et al. (2017) and J18, including the specific likelihood equations and the values for the priors on each parameter. The code is publicly available at https://github.com/djones1040/BEAMS.

### A.2. The BEAMS with Bias Corrections Method

The BBC likelihood has the same form as Equation (4) for SNe Ia but uses Monte Carlo SNANA simulations of the expected CC SN population in the MDS to put a non-Gaussian prior on the $z$-dependent Hubble residuals expected from CC SN contamination. Alternatively, a more flexible, parameterized CC SN model from Hlozek et al. (2012) can be specified within BBC. In contrast to the PSBEAMS method, which also uses a parameterized CC SN model, Hlozek et al. (2012) treat the CC SN distances and dispersion as a polynomial function of the redshift.

Using the BBC method with a simulated CC SN prior could be sensitive to inaccurate CC modeling in the simulations. However, we use both the simulation-based and Hlozek et al. (2012) CC SN likelihood models to ameliorate this concern. We also estimate biases in this method by simulating several

variants of the CC SN population (Section 3.4), each of which differs from the CC SN simulation used to generate the simulation-based prior.

The BBC method has a parameter to renormalize $P(\text{Ia})$ values, effectively the $A$ parameter in Equation (5), but, unlike PSBEAMS, does not have a parameter to linearly shift $P(\text{Ia})$ values. We found in previous work (Jones et al. 2017) that PSNID is an excellent classifier but classifies a significant fraction of CC SNe as being of Type Ia with 100% confidence. For this reason, allowing a free parameter that linearly shifts probabilities up/down (but restricting to $0 < P(\text{Ia}) < 1$) is a necessary parameter when using PSNID probabilities (Jones et al. 2017). Therefore, we use the NN classifier as our baseline BBC classifier but include the use of PSNID probabilities in the systematic uncertainty budget.

We test the BBC method to ensure that it is as reliable as our nominal method (which was validated in Jones et al. 2017). We simulate photometrically classified SN samples and investigate the change in measured SN Ia distances when CC SNe are included in the sample versus when they are excluded. This comparison gives the expected bias in final SN Ia distances that could be caused by inaccurate prior probabilities or imperfections in the likelihood model. To ensure that our test sample is as close as possible to the real data, we combine the set of real, spectroscopically classified SNe Ia used in this analysis with 25 simulated samples of photometrically classified SNe. Each of the 25 simulated samples contains the same number of photometrically classified SNe as our real data; however, unlike the real photometrically classified data, we know the true types of every SN. For each of the 25 samples, we estimate the bias in measured distance by comparing the BBC results from a sample *with* simulated CC SNe to that same sample after excluding simulated CC SNe. The bias is $\lesssim 1$ mmag across the entire redshift range (Figure 7; "default" simulation). However, we caution that the method of estimating the probabilities is not entirely independent of the BBC method; each SN Ia uses the same SALT2 simulation parameters as the Monte Carlo simulation used to generate the data. For this reason, we test additional classification methods and simulations (Section 3.2), which have increased systematic shifts.

## ORCID iDs

D. O. Jones https://orcid.org/0000-0002-6230-0151
R. Kessler https://orcid.org/0000-0003-3221-0419
K. C. Chambers https://orcid.org/0000-0001-6965-7789
D. A. Coulter https://orcid.org/0000-0003-4263-2228
M. E. Huber https://orcid.org/0000-0003-1059-9603
S. W. Jha https://orcid.org/0000-0001-8738-6011
C. D. Kilpatrick https://orcid.org/0000-0002-5740-7747
G. Narayan https://orcid.org/0000-0001-6022-0484
E. Berger https://orcid.org/0000-0002-9392-9681
H. Flewelling https://orcid.org/0000-0002-1050-4056
E. A. Magnier https://orcid.org/0000-0002-7965-2815
S. J. Smartt https://orcid.org/0000-0002-8229-1731
R. J. Wainscoat https://orcid.org/0000-0002-1341-0952
C. Waters https://orcid.org/0000-0003-1989-4879

## References

Alam, S., Ata, M., Bailey, S., et al. 2017, MNRAS, 470, 2617
Amanullah, R., Johansson, J., Goobar, A., et al. 2015, MNRAS, 453, 3300
Anderson, L., Aubourg, É., Bailey, S., et al. 2014, MNRAS, 441, 24
Arnouts, S., & Ilbert, O. 2011, LePHARE: Photometric Analysis for Redshift Estimate, Astrophysics Source Code Library, ascl:1108.009
Astier, P., Guy, J., Regnault, N., et al. 2006, A&A, 447, 31
Becker, A. 2015, HOTPANTS: High Order Transform of PSF ANd Template Subtraction, Astrophysics Source Code Library, ascl:1504.004
Bertin, E., & Arnouts, S. 1996, A&AS, 117, 393
Betoule, M., Kessler, R., Guy, J., et al. 2014, A&A, 568, A22
Bohlin, R. C. 2014, arXiv:1403.6861
Brout, D., Scolnic, D., Kessler, R., et al. 2019, ApJ, 874, 150
Brown, J. S., Stanek, K. Z., Holoien, T. W.-S., et al. 2019, MNRAS, 484, 3785
Bruzual, G., & Charlot, S. 2003, MNRAS, 344, 1000
Bulla, M., Goobar, A., & Dhawan, S. 2018, MNRAS, 479, 3663
Burns, C. R., Parent, E., Phillips, M. M., et al. 2018, ApJ, 869, 56
Burns, C. R., Stritzinger, M., Phillips, M. M., et al. 2014, ApJ, 789, 32
Chambers, K. C., Magnier, E. A., Metcalfe, N., et al. 2016, arXiv:1612.05560
Chevallier, M., & Polarski, D. 2001, IJMPD, 10, 213
Childress, M. J., Wolf, C., & Zahid, H. J. 2014, MNRAS, 445, 1898
Chotard, N., Gangler, E., Aldering, G., et al. 2011, A&A, 529, L4
Conley, A., Guy, J., Sullivan, M., et al. 2011, ApJS, 192, 1
Contreras, C., Hamuy, M., Phillips, M. M., et al. 2010, AJ, 139, 519
Dai, M., & Wang, Y. 2016, MNRAS, 459, 1819
DES Collaboration, Abbott, T. M. C., Allam, S., et al. 2019, ApJL, 872, L30
Filippenko, A. V., Li, W. D., Treffers, R. R., & Modjaz, M. 2001, in ASP Conf. Ser. 246, IAU Coll. 183, Small Telescope Astronomy on Global Scales, ed. B. Paczynski, W.-P. Chen, & C. Lemme (San Francisco, CA: ASP), 121
Flewelling, H. A., Magnier, E. A., Chambers, K. C., et al. 2016, arXiv:1612.05243
Folatelli, G., Phillips, M. M., Burns, C. R., et al. 2010, AJ, 139, 120
Foley, R. J., & Mandel, K. 2013, ApJ, 778, 167
Foley, R. J., Scolnic, D., Rest, A., et al. 2018, MNRAS, 475, 193
Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013, PASP, 125, 306
Gaia Collaboration, Prusti, T., de Bruijne, J. H. J., et al. 2016, A&A, 595, A1
Garnavich, P. M., Jha, S., Challis, P., et al. 1998, ApJ, 509, 74
Goobar, A., Dhawan, S., & Scolnic, D. 2018, MNRAS, 477, L75
Guillochon, J., Parrent, J., Kelley, L. Z., & Margutti, R. 2017, ApJ, 835, 64
Gupta, R. R., Kuhlmann, S., Kovacs, E., et al. 2016, AJ, 152, 154
Guy, J., Astier, P., Baumont, S., et al. 2007, A&A, 466, 11
Guy, J., Sullivan, M., Conley, A., et al. 2010, A&A, 523, A7
Hayden, B. T., Gupta, R. R., Garnavich, P. M., et al. 2013, ApJ, 764, 191
Hicken, M., Challis, P., Jha, S., et al. 2009a, ApJ, 700, 331
Hicken, M., Challis, P., Kirshner, R. P., et al. 2012, ApJS, 200, 12
Hicken, M., Wood-Vasey, W. M., Blondin, S., et al. 2009b, ApJ, 700, 1097
Hill, R., Shariff, H., Trotta, R., et al. 2018, MNRAS, 481, 2766
Hlozek, R., Kunz, M., Bassett, B., et al. 2012, ApJ, 752, 79
Holoien, T. W.-S., Brown, J. S., Stanek, K. Z., et al. 2017, MNRAS, 471, 4966
Hounsell, R., Scolnic, D., Foley, R. J., et al. 2018, ApJ, 867, 23
Howell, D. A. 2001, ApJL, 554, L193
Huber, M., Chambers, K. C., Flewelling, H., et al. 2015, ATel, 7153
Jha, S., Kirshner, R. P., Challis, P., et al. 2006, AJ, 131, 527
Jones, D. O., Riess, A. G., Scolnic, D. M., et al. 2018a, ApJ, 867, 108
Jones, D. O., Scolnic, D. M., Riess, A. G., et al. 2017, ApJ, 843, 6
Jones, D. O., Scolnic, D. M., Riess, A. G., et al. 2018b, ApJ, 857, 51
Jones, D. O., Scolnic, D. M., & Rodney, S. A. 2015, PythonPhot: Simple DAOPHOT-type Photometry in Python, Astrophysics Source Code Library, ascl:1501.010
Jönsson, J., Sullivan, M., Hook, I., et al. 2010, MNRAS, 405, 535
Kelly, P. L., Hicken, M., Burke, D. L., Mandel, K. S., & Kirshner, R. P. 2010, ApJ, 715, 743
Kessler, R., Becker, A. C., Cinabro, D., et al. 2009, ApJS, 185, 32
Kessler, R., Brout, D., D'Andrea, C. B., et al. 2019, MNRAS, 485, 1171
Kessler, R., Guy, J., Marriner, J., et al. 2013, ApJ, 764, 48
Kessler, R., Marriner, J., Childress, M., et al. 2015, AJ, 150, 172
Kessler, R., & Scolnic, D. 2017, ApJ, 836, 56
Kim, Y.-L., Smith, M., Sullivan, M., & Lee, Y.-W. 2018, ApJ, 854, 24
Knights, M., Bassett, B. A., Varughese, M., et al. 2013, JCAP, 1, 039
Knop, R. A., Aldering, G., Amanullah, R., et al. 2003, ApJ, 598, 102
Kowalski, M., Rubin, D., Aldering, G., et al. 2008, ApJ, 686, 749
Kunz, M., Bassett, B. A., & Hlozek, R. A. 2007, PhRvD, 75, 103508
Lampeitl, H., Smith, M., Nichol, R. C., et al. 2010, ApJ, 722, 566
Lewis, A., & Bridle, S. 2002, PhRvD, 66, 103511
Li, W., Leaman, J., Chornock, R., et al. 2011, MNRAS, 412, 1441
Linder, E. V. 2003, PhRvL, 90, 091301
Mandel, K. S., Narayan, G., & Kirshner, R. P. 2011, ApJ, 731, 120
Marriner, J., Bernstein, J. P., Kessler, R., et al. 2011, ApJ, 740, 72

Ménard, B., Scranton, R., Fukugita, M., & Richards, G. 2010, MNRAS, 405, 1025
Perlmutter, S., Aldering, G., Goldhaber, G., et al. 1999, ApJ, 517, 565
Pierel, J. D. R., Rodney, S., Avelino, A., et al. 2018, PASP, 130, 114504
Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2016, A&A, 594, A13
Planck Collaboration, Aghanim, N., Akrami, Y., et al. 2018, arXiv:1807.06209
Rest, A., Scolnic, D., Foley, R. J., et al. 2014, ApJ, 795, 44
Rest, A., Stubbs, C., Becker, A. C., et al. 2005, ApJ, 634, 1103
Riess, A. G., Casertano, S., Yuan, W., et al. 2018, ApJ, 855, 136
Riess, A. G., Filippenko, A. V., Challis, P., et al. 1998, AJ, 116, 1009
Riess, A. G., Kirshner, R. P., Schmidt, B. P., et al. 1999, AJ, 117, 707
Riess, A. G., Strolger, L.-G., Casertano, S., et al. 2007, ApJ, 659, 98
Riess, A. G., Strolger, L.-G., Tonry, J., et al. 2004, ApJ, 607, 665
Rigault, M., Aldering, G., Kowalski, M., et al. 2015, ApJ, 802, 20
Rigault, M., Brinnel, V., Aldering, G., et al. 2018, A&A, submitted (arXiv:1806.03849)
Rigault, M., Copin, Y., Aldering, G., et al. 2013, A&A, 560, A66
Roberts, E., Lochner, M., Fonseca, J., et al. 2017, JCAP, 10, 036
Roman, M., Hardin, D., Betoule, M., et al. 2018, A&A, 615, A68
Ross, A. J., Samushia, L., Howlett, C., et al. 2015, MNRAS, 449, 835
Sako, M., Bassett, B., Becker, A. C., et al. 2018, PASP, 130, 064002
Sako, M., Bassett, B., Connolly, B., et al. 2011, ApJ, 738, 162
Schechter, P. L., Mateo, M., & Saha, A. 1993, PASP, 105, 1342
Schlafly, E. F., Finkbeiner, D. P., Jurić, M., et al. 2012, ApJ, 756, 158
Scolnic, D., Casertano, S., Riess, A., et al. 2015, ApJ, 815, 117
Scolnic, D., & Kessler, R. 2016, ApJL, 822, L35
Scolnic, D., Rest, A., Riess, A., et al. 2014a, ApJ, 795, 45
Scolnic, D. M., Jones, D. O., Rest, A., et al. 2018, ApJ, 859, 101
Scolnic, D. M., Riess, A. G., Foley, R. J., et al. 2014b, ApJ, 780, 37
Spergel, D., Gehrels, N., Baltay, C., et al. 2015, arXiv:1503.03757
Stetson, P. B. 1987, PASP, 99, 191
Stritzinger, M. D., Phillips, M. M., Boldt, L. N., et al. 2011, AJ, 142, 156
Sullivan, M., Conley, A., Howell, D. A., et al. 2010, MNRAS, 406, 782
Sullivan, M., Guy, J., Conley, A., et al. 2011, ApJ, 737, 102
Sullivan, M., Le Borgne, D., Pritchet, C. J., et al. 2006, ApJ, 648, 868
Tartaglia, L., Sand, D. J., Valenti, S., et al. 2018, ApJ, 853, 62
The LSST Dark Energy Science Collaboration, Mandelbaum, R., Eifler, T., et al. 2018, arXiv:1809.01669
Tonry, J. L., Denneau, L., Heinze, A. N., et al. 2018, PASP, 130, 064505
Tonry, J. L., Schmidt, B. P., Barris, B., et al. 2003, ApJ, 594, 1
Tripp, R. 1998, A&A, 331, 815
Wood-Vasey, W. M., Miknaitis, G., Stubbs, C. W., et al. 2007, ApJ, 666, 694