The California Legacy Survey: A Three-Decade Census of Extrasolar Planets

Thesis by
Lee J. Rosenthal

In Partial Fulfillment of the Requirements for the
Degree of
Doctor of Philosophy

Caltech

CALIFORNIA INSTITUTE OF TECHNOLOGY
Pasadena, California

2022
Defended April 29th, 2022

© 2022

Lee J. Rosenthal
ORCID: 0000-0001-8391-5182

All rights reserved

ACKNOWLEDGEMENTS

In the words of an unattributed proverb / Al Gore / the 30 Rock writer's room: If you want to go fast, go alone. If you want to go far, go together. I am leaving Caltech after traveling far along unexpected paths, and am immensely grateful to the many people who helped me along the way.

Thank you to Andrew Howard, who took a chance on me when I wanted to try something new. I have trouble imagining a more thoughtful and patient advisor, willing to consider new ideas without bias and walk step by step through any challenge. Andrew has been a model for ease and rational thinking in the face of uncertainty, both in science and in life. I will always appreciate his particular brand of calm kindness.

Thank you to Heather Knutson, who helped me find my way into a new field and provided both professional and personal support when I needed it most. I often walked into her office anxious and walked out feeling comforted, bolstered, and excited by new ideas that she would casually propose. Heather is a deeply generous mentor and has made me a better statistician, writer, and person.

Thanks also to the rest of my thesis committee. Lynne Hillenbrand has made me a sharper scientist and nudged me to focus on what matters most and speak my mind. Every conversation I ever had with Konstantin Batygin brought curiosity and joy, whether it was about giant planets, obscure genres of music, or bizarre incidents at Caltech (sometimes all at once). Gregg Hallinan was an excellent advisor to me during my first two years in graduate school, and I appreciate his support of my desire to branch out both in and beyond our work together.

I am lucky to have the help and friendship of BJ Fulton, my academic older brother. BJ helped me throughout my first year in the California Planet Search team and beyond, with coding assistance and assessment of research projects and more. Thanks also to the rest of the California Planet Search team for friendship and scientific collaboration.

I have many academic mentors to thank from before my time at Caltech. Volker Springel was a gracious host to me during my Fulbright year in Heidelberg, and I would never have made it there without the guidance of Howard Smith and Chris Hayward during the previous summer. I am grateful to the professors at Haverford College who taught me and guided me on the path to graduate school. Peter Love taught my first physics course with zest and style, and sent me to Strawbridge Observatory to learn from Steve Boughn. Steve and his order-of-magnitude seminar immediately hooked me on astrophysics, leading me to the tutelage of the great researcher and teacher Desika Narayanan. I would not be graduating from Caltech without the steady hand of my academic advisor Beth Willman, who has been there for me during many moments of growth and hardship. Before all these people came my high school math teacher Michael Londino, who taught me that math can be simple and that persistence is everything.

I am thankful to the Caltech astrophysics graduate students for continuing the kind environment that I found when I first visited the department. Thanks particularly to Anna Ho, my mentor and friend, for sharing her wisdom and penchant for adventures, from long hikes to starring roles at the Ice House.

It's hard to explain how important my friends at Caltech and across Los Angeles have been to me during the past six years. Thank you to Sophia Charan, Jacob Wasserman, Sophie Miller, and Alex Welch for good times watching The Good Place. Thank you to Noah Olsman, Scott Barenfeld, and Donal O'Sullivan for many laughs and tidbits of advice. Thank you to Heidi Klumpe, Matteo Ruggero Ronchi, Milan Cvitcovic, and many others for teaching me to always 'say yes and'. Playing with Caltech's improv group gave me some of my fondest memories from graduate school. Thank you to Michael Lewis and Jess Solomon for kindness and reminders that there is life outside of Caltech.

One of the unexpected delights of the past two years has been the growth of friendships that will last beyond my PhD. Thanks to Hal Baron for many bike rides through Pasadena, to Avi Flamholz and Simon Mahler for post-minyan breakfast chats, to David and Aviva Zierler for making me feel like part of their family, and to Hershy and Sheva Stolik for many hours of singing, study, and celebration. Thank you to the book club crew for keeping life interesting: Zachary Sercel, Oren Mizrahi, Michael Wolman, Helen Wexler, and Noa Wexler, the best of all of us.

I am grateful, grateful, grateful for the friends I have made and kept since before my time at Caltech. Belinda Schwartz, Dan Listwa, Josh Mermelstein, Matthew Sirkin, and Warren Brodsky have shown me what it means to stay close with someone when a continent separates you. I look forward to sharing many celebrations throughout our lives. Thanks also to Harry Levin, the best Honor Code Officer anyone could ask for and a testament to both kindness and grit.

A special thanks to my great aunt and uncle Isa and Bill Aron. They brought me into the Los Angeles Jewish community with loving arms and provided a home away from home when I needed it most. I will miss sharing their company, crashing in their spare room over Shabbat, and eating all their food.

Another special thanks to Joyce Raynor, Howard Tepper, and my great uncle Jimmy Schwarz. Being in a study group with them is another surprise joy from the past two years. “Who is wise? One who learns from all people” (Pirkei Avot 4:1). I feel very lucky to learn from and with this group, and hopefully have become a better person in the process.

One final special thanks to my grandmother Adele Greenspun, and my late grandparents Steven Schwarz, Burt Greenspun, Pat Rosenthal, and Marty Rosenthal z"l. They each have shown me a different side of myself and given me more love than I could ever ask for.

My parents Don and Joanie Rosenthal have been there for me every step of the way, in this and every other phase of my life. They are my first and last inspirations to be kind, thoughtful, and decisive about doing the right thing. Any creativity, leadership, or compassion I might express is owed to them. My brother Dean Rosenthal has inspired me to persevere through uncertainty, enjoy life, and take myself less seriously. I am quite excited to share some time and space with him in Boston.

I am having trouble putting into words what it means to feel so supported by so many people, so I will just end by saying thank you again and I appreciate you all.

ABSTRACT

Taking an accurate census of planets orbiting other stars, otherwise known as exoplanets, is a crucial step toward understanding the nature of planet formation and placing Earth and the Solar System in a broader context. For my thesis, I present the culmination of a three-decade a high-precision radial velocity (RV) survey of 719 FGKM stars, known as the California Legacy Survey (CLS), and use this survey to perform a statistical study of exoplanets. I developed computational methods for planet search and RV sensitivity characterization, and detected 164 known exoplanets and 14 newly discovered or revised exoplanets and substellar companions in the CLS. I used this star and planet catalog to measure the occurrence rates of several exoplanet subtypes and probe formation pathways. I found that giant planet occurrence is greatly enhanced beyond 1 Earth-Sun distance (au), then decreases beyond 8 au. This implies that giant planet formation is much easier beyond the water-ice line of most stars, possibly due to greater ease of solid coagulation and pebble accretion, and eventually decreases with the density profile of of protoplanetary disks. It also means that Jupiter and Saturn are located in the orbital space of greatest giant planet occurrence, making the Solar System typical in giant placement. I then investigated the relationship between small close-in planets and cold outer giants, and found that not all giants host inner rocky companions. Rather, giants under about one-third of a Jupiter mass or beyond ~ 3 au are more likely to have small companions than their warmer and more massive counterparts. Finally, I compared single giant and multi giant planetary systems. I performed novel characterizations of the orbital eccentricity distributions of these two populations, and discovered that multi-giant stellar hosts are on average significantly more metal-rich than single-giant hosts. This means that the Sun with its two giant planets is atypical among much more metal-rich giant hosts. I also found that lonely giants present a pile-up of 'super-hot Jupiters' within 0.06 au not shared by neighborly giants, and that giants orbiting the same star tend to have similar masses. Taken together, these findings are a substantial leap forward in understanding the architectures and origins of planetary systems.

PUBLISHED CONTENT AND CONTRIBUTIONS

Fulton, B. J. et al. (July 2021). “California Legacy Survey. II. Occurrence of Giant Planets beyond the Ice Line”. In: *ApJS* 255.1, 14, p. 14. DOI: 10.3847/1538-4365/abfcc1. arXiv: 2105.11584 [astro-ph.EP].
I wrote most of the document, wrote the analysis code used to produce all results, and helped with figure construction.

Rosenthal, L. J., B. J. Fulton, et al. (July 2021). “The California Legacy Survey. I. A Catalog of 178 Planets from Precision Radial Velocity Monitoring of 719 Nearby Stars over Three Decades”. In: *ApJS* 255.1, 8, p. 8. DOI: 10.3847/1538-4365/abe23c. arXiv: 2105.11583 [astro-ph.EP].
I was the primary author, and performed all writing, analysis, and production in the paper.

Rosenthal, L. J., H. A. Knutson, et al. (Dec. 2021). “The California Legacy Survey III. On The Shoulders of (Some) Giants: The Relationship between Inner Small Planets and Outer Massive Planets”. In: *arXiv e-prints*, arXiv:2112.03399, arXiv:2112.03399. arXiv: 2112.03399 [astro-ph.EP].
I was the primary author, and performed all writing, analysis, and production in the paper.

TABLE OF CONTENTS

Acknowledgements iii
Abstract vi
Published Content and Contributions vii
Table of Contents vii
List of Illustrations x
List of Tables xxiv
Chapter I: Introduction 1
1.1 Is The Solar System 'Typical'? 1
1.2 The Nature of Planets 2
1.3 Beyond Sol 3
1.4 Origin Story 5
1.5 Counting to Answer Questions 6
1.6 Taking a New Census 9
Chapter II: Taking a Census: A Catalog of Planets Discovered From Radial Velocity Monitoring 11
2.1 Introduction 11
2.2 Stellar Sample Selection 13
2.3 Observations 15
2.4 Stellar properties 20
2.5 Planet Catalog Methods 21
2.6 Planet and Stellar/Substellar Companion Catalog 34
2.7 Discussion 35
2.8 Supplemental Information 38
Chapter III: Frost Giants: The Occurrence of Jovian Planets Beyond the Ice line 83
3.1 Introduction 83
3.2 Survey Review 84
3.3 Methods 86
3.4 Results 89
3.5 Discussion 94
3.6 Conclusion 101
Chapter IV: On The Shoulders of (Some) Giants: The Relationship Between Inner Small Planets and Outer Massive Planets 103
4.1 Introduction 103
4.2 Survey Review 105
4.3 Methods 107
4.4 Results 110
4.5 Discussion 115
4.6 Conclusion 124

Chapter V: Lonely, Poor, And Eccentric: A Comparison Between Solitary And Neighborly Gas Giants . . . 125
5.1 Introduction . . . 125
5.2 Survey Review . . . 126
5.3 Methods . . . 128
5.4 Results . . . 130
5.5 Discussion . . . 142
5.6 Conclusions . . . 146
Chapter VI: Conclusions . . . 148
Bibliography . . . 151

LIST OF ILLUSTRATIONS

Number *Page*

2.1 Venn diagram showing the overlap between the stars in the Keck Planet Search sample (Cumming et al., 2008), the Eta-Earth sample (Howard, Johnson, Marcy, Fischer, Wright, Bernat, et al., 2010), and a 25 pc northern hemisphere volume-limited survey (Hirsch et al. in prep). 31 stars in the California Legacy Survey do not belong to the union of these three surveys. 16
2.2 Distributions of observational baseline versus number of observations. The top left panel shows these statistics for all stars in the CLS sample; the top right panel shows stars around which we detect planets; the bottom panel shows stars around which we do not detect planets. Median baseline and number of observations for each sample are overplotted as translucent circles. 19
2.3 Stellar property measurements of the California Legacy Survey, in effective temperature, surface gravity, and mass. The sample consists of stars spanning spectral types F, G, K, and M, some of which have evolved off of the main sequence. Most stars have metallicities within 0.4 dex of Solar metallicity, with the exception of a small handful of extremely metal-poor stars, which lie below the main sequence on this plot. 21
2.4 Stellar parameter distributions. The left column shows mass, metallicity, and effective temperature, while the right column shows parallax-inferred distance, V-band magnitude, and $B - V$ color. Black lines are histograms of the stellar parameter median values. For the left column, colored lines are 500 histograms per panel, with parameters redrawn from normal distributions with width equal to their individual measurement uncertainties. We left these redrawn parameter histograms out for the plots in the right column because distance, magnitude, and color have uncertainties that are smaller than the chosen bin size. 22
2.5 Search algorithm flowchart. 26

2.6 Example `RVSearch` summary plot, for the known two-planet system HIP 109388. Panel a) shows the total model plotted over the radial velocity time series, while panel b) shows the model residuals. Each successive row shows a phase-folded signal discovered by `RVSearch` on the left, and the associated periodogram on the right. The final row shows the running periodograms of each signal, generated with Lomb–Scargle power, on the left, and the final periodogram on the right. 27
2.7 RVs and completeness contours for three datasets with similar baselines, median measurement errors, and stellar jitter. The left column plots RVs with respect to time, while the right column plots injected signals in the $M \sin i$ and a plane, where blue dots are recovered injections and red dots are not. The right column also shows detection probability contours, with 50% plotted as a solid black line. From top to bottom, we show RVs and contours for HD 44420, for which we have 24 RVs; HD 97343, for which we have 94 RVs; and HD 12051, for which we have 372 RVs. 29
2.8 Stack of all final nondetection periodograms in the CLS planet search, linearly interpolated to the same period grid. A broad peak around 1 year is evident, as well as narrow peaks at 1/2-year, 1/3-year, and 1/4-year. 32
2.9 Candidate vetting flowchart. 33
2.10 Scatterplot of best-fit $M \sin i$ and semi-major axis values for planets in the CLS catalog. Blue dots represent known planets, while green circles represent newly discovered planets and planets with significantly revised orbits. 34
2.11 Contours (1- and 2-σ) of $M \sin i$ and semi-major axes for planets in the CLS sample whose semi-major axis posteriors extend beyond 10 AU. Contours for HD 26161 b have hard cutoffs due to sparsity below 7 M_J and 12 AU; these limits come from the data's baseline and RV increase to date. 36

2.12 Orbit gallery for six of the coldest companions in our survey. We plot RV data and Keplerian model versus year, and subtract off the model signatures of inner companions and stellar activity. We did not include UMa 47 d, seen in Figure 2.11, in this plot, because its detection relied on early Lick-Hamilton RVs, and we wanted to showcase HIRES RV measurements from the past twenty-four years. . 37
2.13 $M \sin i$, a, and eccentricity of the CLS sample. Eccentricity is plotted in medians and 68% confidence intervals, while scatter size is proportional to $M \sin i$ posterior mode. 38
2.14 `RVSearch` summary plot for HD 3765. See Figure 2.6 for plot description. 39
2.15 Lomb-Scargle periodogram of HIRES S-values for HD 3765. Significant power at and beyond 4,300 days. 40
2.16 `RadVel` model orbital plot for HD 3765, including a Gaussian process with a squared-exponential kernel. The grey shaded curve represents the 68% interval for the Gaussian process RV signature. 41
2.17 Orbital and Gaussian process parameter posteriors for HD 3765. η_1 is the GP amplitude, while η_2 is the GP exponential decay timescale. 42
2.18 `RVSearch` summary plot for HD 24040. See Figure 2.6 for plot description. 43
2.19 Lomb-Scargle periodogram of HIRES S-values for HD 24040. No periods show power that is statistically significant. 44
2.20 `RVSearch` summary plot for HD 26161. See Figure 2.6 for plot description. 45
2.21 Rejection sampling posterior for HD 26161. 46
2.22 Possible orbits for HD 26161 b. RV curves are drawn from the rejection sampling posterior generated with `TheJoker`. The color of each orbit drawn from the posterior scales with $M \sin i$ 47
2.23 `RVSearch` summary plot for HD 66428. See Figure 2.6 for plot description. 48
2.24 `RVSearch` summary plot for HD 68988. See Figure 2.6 for plot description. 49
2.25 `RVSearch` summary plot for HD 95735. See Figure 2.6 for plot description. 51

2.26 Lomb-Scargle periodogram of HIRES S-values for HD 95735. There is evidence for an activity cycle longer than 10,000 days, but no significant power near the period of our 3,000-day planet candidate. . . 52
2.27 `RVSearch` summary plot for HD 107148. See Figure 2.6 for plot description. 53
2.28 Lomb-Scargle periodogram of HIRES S-values for HD 107148. Significant power at and beyond 4,300 days. 54
2.29 `RVSearch` summary plot for HD 136925. See Figure 2.6 for plot description. 55
2.30 Periodogram of HIRES S-values for HD 136925. Significant periodicity around 300 days, near the period of the inner signal. 56
2.31 `RVSearch` summary plot for HD 141004. See Figure 2.6 for plot description. 57
2.32 `RVSearch` summary plot for HD 145675. See Figure 2.6 for plot description. 58
2.33 Lomb-Scargle periodogram of HIRES S-values for HD 145675 showing significant power at 3,600 days. 59
2.34 Activity vetting plots for HD 145675. For all panels, the horizontal axis shows the S-value activity metric of each observation, while the vertical axis shows corresponding RV residuals for each individual Keplerian orbit. The left-hand panels show HIRES post-upgrade observations, while the right-hand panels show APF observations. Each row shows RVs with the model residuals of one Keplerian model, with the other Keplerian models subtracted from the data. The blue lines show linear correlations between these residuals and the corresponding S-values. In the HIRES and APF data, we measured $> 3\sigma$ correlations for the third Keplerian signal. The APF and HIRES linear correlations are within 3σ of each other, implying that this signal is caused by stellar activity. We find correlations between the residuals and S-values for the second signal as well, but they are significantly different for HIRES and APF. Since the period of this signal is much greater than the APF baseline of this star, we discount this second correlation as caused by the limited baseline of the data with respect to the signal. 60
2.35 `RVSearch` summary plot for HD 156668. See Figure 2.6 for plot description. 61

2.36 RVSearch summary plot for HD 164922. See Figure 2.6 for plot description. 63
2.37 Visualization of APT photometry analysis for HD 164922. The top panel shows a Lomb-Scargle periodogram of the photometry, with a moving-average filter to reduce alias issues. The middle panel shows an unfiltered periodogram. 64
2.38 Visualization of HIRES S-value analysis for HD 164922. The top panel shows a Lomb-Scargle periodogram of the S-values, with a moving-average filter to reduce alias issues. The middle panel shows an unfiltered periodogram. 65
2.39 RVSearch summary plot for HD 168009. See Figure 2.6 for plot description. 66
2.40 RVSearch summary plot for HD 213472. See Figure 2.6 for plot description. 67
2.41 Rejection sampling posterior for HD 213472 b orbital parameters. $\Delta\gamma$ is the relative linear offset between different instrumental datasets, in this case pre-upgrade and post-upgrade HIRES. 68
2.42 Possible orbits for HD 213472 b. RV curves are drawn from the rejection sampling posterior generated with TheJoker. The color of each orbit drawn from the posterior scales with $M \sin i$ 69
2.43 RVSearch summary plot for HD 115617. See Figure 2.6 for plot description. Note the nearly equivalent-height peaks at 1/3 and 1/4 year in panel h, corresponding to the 124 day reported planet. Panel j shows that there is residual power at 1 year after subtracting the 122 day signal, suggesting the presence of yearly systematic noise in the data. 71
2.44 PSF Lomb–Scargle periodograms for HD 115617. Each panel corresponds to a Doppler code PSF fitting parameter. 72
2.45 PSF correlation plots for the candidate HD 115617 d. Each panel corresponds to a Doppler code PSF fitting parameter, with PSF value on the x-axis and RV without the signatures of the inner two planets on the y-axis. Dashed blue lines are least-squares linear fits. R is the Pearson correlation value; multiple PSF parameters have $|R| > 0.15$. . . 73

2.46 HIRES post-upgrade RV and S-value activity timeseries for HD 154345. Note that the two datasets share minima and appear to be in phase when post-upgrade observations began, but have drifted completely out of phase over the following 23 years. 74

2.47 `RVSearch` summary plot for HD 154345; see Figure 2.6 for description. `RVSearch` first recovered a strong signal at 9 years, but then recovered additional power at a similar period due to stellar activity. The final orbit fit switched the two models, so that panels e) and d) show the planetary signal, while panels c) and f) show the stellar activity signal. 75

2.48 Stellar rotation analysis of HIRES S-values for HD 26965. The top panel shows a Lomb–Scargle periodogram of the S-values after we applied a high-pass filter to them, to remove the impact of the long-period magnetic activity cycle. The middle panel shows a periodogram of the raw S-values. The top panel shows significant periodicity near 40 days, with a maximum at 41.6 days. The bottom-left panel shows the filtered S-values, while the bottom-right panel shows the filtered S-values phased to 41.6 days; there appears to be a coherent signal at this period, implying stellar rotation with this period. 76

2.49 HIRES post-upgrade S-values and RVs for HD 26965. The two datasets both have long-period power and are in phase with each other. 77

2.50 `RVSearch` summary plot for HD 34445; see Figure 2.6 for description. `RVSearch` first recovered the known giant planet, then a series of what are likely spurious signals caused by yearly aliasing or stellar activity. 78

2.51 HIRES post-upgrade RV and S-value activity timeseries for HD 34445, with the giant planet RV model subtracted. Note that these two datasets share a negative long-term trend, which we believe accounts for the claimed 5,700-day planet in the system. 79

2.52 Stellar rotation analysis of HIRES S-values for HD 34445. The top panel shows a Lomb–Scargle periodogram of the S-values after we applied a high-pass filter to them, to remove the impact of the long-period magnetic activity cycle. The middle panel shows a periodogram of the raw S-values. The top panel shows significant periodicity around 52.1 days. The bottom-left panel shows the filtered S-values, while the bottom-right panel shows the filtered S-values phased to 52.1 days; there appears to be a coherent signal at this period, implying stellar rotation with this period. This led us to label the 49 day claimed planet as a false positive, since there is insufficient evidence to distinguish it from stellar rotation. 80

2.53 PSF Lomb–Scargle periodograms for HD 34445. Each panel corresponds to a Doppler code PSF fitting parameter. 81

2.54 PSF correlation plots for HD 34445, without the RV signature of the star's giant planet. Each panel corresponds to a Doppler code PSF fitting parameter, with PSF value on the x-axis and RV without the giant planet signature on the y-axis. Dashed blue lines are least-squares linear fits. R is the Pearson correlation value; multiple PSF parameters have $|R| > 0.15$. 82

3.1 California Legacy Survey planet catalog and survey-averaged search completeness contours in semi-major axis and $M \sin i$. 3% and 1% search completeness contours are highlighted in white. 86

3.2 Non-parametric occurrence rates for semi-major axes of 0.03–30 au for planets with minimum masses from 30–6000 $M \sin i$ assuming uniform occurrence across $\ln(M \sin i)$. The dashed blue line represents a planet count in each semi-major axis bin without correcting for completeness; bold lines and dots show the maximum posterior values for the Poisson likelihood model; vertical lines represent 15.9–84.1% confidence intervals (except for the last bin, which is not separated from zero and shows 0–68.2%); and transparent steps show draws from the occurrence posterior. We see a clear enhancement around 1–10 au, and a tentative falloff beyond that range. 89

3.3 Our broken power law model, juxtaposed with our non-parametric model and measurements from Fernandes et al., 2019 and Wittenmyer, Wang, et al., 2020. The transparent curves represent draws from the broken power law posterior. We find that the power law index beyond the break is ~2.5σ-separated from zero, implying an occurrence falloff beyond the water-ice line. Cumming et al. (2008) performed a power-law fit to the occurrence rates of planets orbiting only within 3 au; the light dotted blue line represents an extrapolation to wider separations. 90
3.4 Broken power law posterior. C is a normalization constant, β is the power law index beyond the break, a_0 determines the location of the break in units of au, and $\beta+\gamma$ is the power law index within the break. The index beyond the break β is ~ 99.1%-separated from zero. 91
3.5 A comparison between sub- and super-Jovian occurrence. Steps and dots show maximum posterior values, and vertical lines show 15.9–84.1% confidence intervals. The sub-Jovians are consistently more common than the super-Jovians, and both populations are enhanced beyond 1 au. Combining these two populations produces the same trends seen when we assume uniform occurrence across all masses. . 93
3.6 Planet occurrence within 1–5 au with respect to $M \sin i$. Steps and dots show maximum posterior values, and vertical lines show 15.9–84.1% confidence intervals. The mass function is constant within 30–900 $M_\oplus$, and falls off beyond 900 $M_\oplus$. 94
3.7 Left: occurrence of giant planets more massive than 100 $M_\oplus$ and within 1–5 au as a function of host star mass, in six splits. Steps and dots show maximum posterior values, and vertical lines show 15.9–84.1% confidence intervals. There is an increase in occurrence beyond roughly 1 $M_\odot$, which is in agreement with Johnson, Aller, et al., 2010's original measurement of giant planet occurrence versus host-star mass. Right: occurrence of giant planets more massive than 100 $M_\oplus$ and within 1–5 au as a function of host star metallicity, in six splits. Steps and dots show maximum posterior values, and vertical lines show 15.9–84.1% confidence intervals. There is a clear increase in occurrence beyond roughly 0.1 dex, which is in agreement with Fischer, Laughlin, et al., 2005's original report of a correlation between giant planet occurrence and host-star metallicity. 95

3.8 Occurrence rate comparison to direct imaging studies. *Left*: Frequency of cool, massive companions with the direct imaging study of Biller et al. (2013). While they did not detect any planets in their survey they were able to put upper limits on the frequency of companions using assumptions of either hot-start (COND) or cold-start (DUSTY) models for planetary formation and infrared brightness. *Right*: Same as left, but compared with the results of Bowler, Liu, et al. (2015) and Nielsen et al. (2019) for the mass and separation limits specified in the x-axis label. The gray shading represents the 95% upper limit on occurrence from Bowler, Liu, et al. (2015). . . . 97

3.9 *Left:* Occurrence rate comparison with the microlensing survey of Cassan et al. (2012). We plot the 1 σ limits from Cassan et al. (2012) as the shaded blue region. The occurrence rate posterior from this work is plotted in black. *Right:* Occurrence rate comparison with the combined analysis of Clanton and Gaudi (2016). The occurrence rate posterior from this work is plotted in black. The 1 σ limits from Clanton and Gaudi (2016) are indicated by the shaded red region. Clanton and Gaudi (2016) combine constraints from direct imaging, microlensing, and previous radial velocity studies. 98

3.10 Analogous to Figure 2, occurrence with respect to stellar light intensity instead of orbital separation. Here we see a similar enhancement in the occurrence rate of giant planets where the insolation flux is equal to that of Earth and tentative evidence for a fall off in occurrence just beyond that. 100

4.1 Minimum mass ($M \sin i$) versus semi-major axis values for small ($M \sin i <$ 30 $M_\oplus$) planets, and their giant outer companions, in the CLS catalog. Diamonds are small planets without outer giants, pentagons are small planets with outer giants, and circles are outer giants. Contours are the completeness map for small planet hosts. . . 107

4.2 A histogram of observed small planet multiplicity in our sample. This is not corrected for search completeness, so it should only be interpreted as the multiplicity of detected planets, not as the underlying multiplicity distribution. There are 719 total stars in the CLS, around 29 of which we have detected small planets. 110

4.3 Left: Two measurements of the conditional occurrence of inner small planets given the presence of an outer gas giant. The black distribution is our direct measurement, while the green distribution uses Bayes Theorem to infer it from other measurements. Right: Our sample of planet pairs with small planets within the region of interest, with our inner small planet box outlined in red and our outer giant box outlined in purple. We assign a number to each planetary system and label individual planets accordingly. 112

4.4 Same as Figure 4.3, but for Jupiter analogs within 3–7 au and 0.3–13 M_J instead of the broader giant population. 112

4.5 Analytical mass distributions for small planets used in this work (green) and from Neil and Rogers, 2020 (purple). This work assumes a uniform distribution in $\ln(M)$. Assuming a uniform distribution in $\ln(a)$, this leads to a 25% recovery rate in our survey of small planets within 2–30 $M_\oplus$ and 0.023–1 au, given our search completeness. Neil and Rogers, 2020 fit a log-normal mixture model to a sample of *Kepler* planets and found a distinct small planet component, shown here. This model leads to an 18.3% recovery rate in our survey. . . . 113

4.6 Left: Occurrence grid for the full CLS sample of 719 stars. Cell shade and number annotation reflect the median expected number of planets per 100 stars in each bin. Empty bins show an expected upper limit on occurrence as the 84.1th percentile on the occurrence rate posterior. Right: Same, but only for the 28 hosts of detected small planets with $M \sin i < 30\ M_\oplus$ and $a < 1$ AU. Note that this sample includes 55 Cnc's four cold giants, whereas our fractional analysis in Figures 4.3 and 4.4 excludes 55 Cnc, since its inner ultra-short period planet is undetectable by our automated search due to our period limits. The right-hand panel shows that there is an absence of warm gas giants in our sample of detected small planet hosts. 113

4.7 Top: cold giant planets without detected inner small companions in the CLS sample, with associated completeness contours. The contours show that the datasets associated with these systems are somewhat sensitive to planets within 0.5 au and 30 $M_{\oplus}$; we have outlined our small-planet parameter space in green for context. This means that we can say with some confidence that not all of these systems harbor undetected small planets. Bottom: Small planets without detected outer companions in the CLS sample, with associated completeness contours. We have outlined our outer giant parameter space in purple for context. 115

4.8 Left: Cumulative metallicity distributions for hosts of lonely small planets and hosts of both inner small and outer giant planets. Solid and dashed steps show distributions of median metallicity measurements, while transparent steps show metallicities drawn many times from Gaussian distributions, with means and standard deviations taken from measurement means and uncertainties. Right: A distribution of P-values from many Kolmogorov-Smirnov tests, performed on 5×10^5 drawn sets of metallicities for the two stellar host groups. 97% of the draws produce $P < 0.05$, and 73% of the draws produce $P < 0.01$. This implies that the underlying metallicity distributions of these two groups are distinct. 116

4.9 A comparison between outer giant companions and the rest of the CLS giant sample in semi-major axis and orbital eccentricity space. Circle size is proportional to $\log(M \sin i)$. Companions to inner small planets are less eccentric than the parent sample, and less massive. . . 118

4.10 Mass functions for all planets and outer companions within 0.23–10 au, produced by marginalizing the occurrence grids in Figure 4.6 along their orbital separation axes. Steps and dots show maximum a posteriori values; vertical bars show 15.9–84.1% confidence intervals. 119

4.11 Observed eccentricity distributions for all giant planets and outer giant companions. This plot shows histograms of the maximum a posteriori eccentricities of our sample. 121

4.12 The results of three Monte Carlo experiments detailed in Subsection 4.5. We estimate how planet multiplicity biases our estimates of the probability that a star hosts at least one small planet, or $P(1+)$. The black vertical line is the true value of $P(1+)$ for all three experiments. The legend shows the multiplicity distribution and resulting average number of planets per host $< N_P >$ for all three experiments. In the one-per-host case, the Poisson model accurately recovers $P(1+)$, while cases with higher $< N_P >$ increasingly overestimate the true value. 123

5.1 Left: Eccentricity with respect to semi-major axis for single giant planets and multiple giant planets. Vertical bars show 15.9–84.1% confidence intervals. Circle radius is proportional to $M \sin i$. Right: Same as left, but color-coded by host star metallicity. Curve shows estimate of tidal disruption limit assuming that periastron at 0.03 au leads to disruption (Dawson and Johnson, 2018). 127

5.2 Planet eccentricity versus host star metallicity for single and multiple giant systems. 127

5.3 Eccentricity distributions of single and multiple giants beyond 0.3 au. Purple/top distribution is for multiple giants, green/bottom distribution is for single giants. Histograms are distributions of individual maximum posterior values, while curves show the median and many draws from beta distribution posteriors, produced using hierarchical inference. 132

5.4 Beta distribution posteriors for the single and multiple giant eccentricity distributions. The two populations are $\sim 3\sigma$-distinct. 133

5.5 Cumulative metallicity distributions for hosts of single giant planets and hosts of multiple giant planets. Solid and dashed steps show distributions of median metallicity measurements, while transparent steps show metallicities drawn many times from Gaussian distributions, with means and standard deviations taken from measurement means and uncertainties. Kolmogorov-Smirnov tests show that the two groups are distinct. 134

5.6 Top: Average search completeness contours for giant hosts below their median metallicity. Bottom: Average search completeness contours for giant hosts above their median metallicity. The two contours are nearly equivalent, implying that sensitivity bias to metal-rich stars is not the cause of the single-multi host distinction seen in Figure 5.5. 135

5.7 Cumulative metallicity distributions for hosts of circular ($e < 0.5$) giant planets, hosts of eccentric ($e \geq 0.5$) giant planets, and multi-giant hosts. See Figure 5.5 for description. The eccentric single hosts are inconclusively drawn from the same underlying metallicity distribution as the less eccentric singles or the multis. 136

5.8 Occurrence with respect to semi-major axis for single giant planets and multiple giant planets. Steps and dots are median values, while vertical lines are 15.9–84.1% confidence intervals. The single giant planets display a pile-up of very hot Jupiters within 0.06 au, tentatively not shared by the multiple giant planets. 137

5.9 Lonely and neighborly warm Jupiters, described in Subsection 5.4. 11 systems have lonely warm Jupiters, while 10 systems host both a warm Jupiter and at least one detected outer giant companion. This does not refute the synthetically driven claim by Jackson et al., 2021 of secular warm Jupiter migration. 138

5.10 Occurrence as a function of $M \sin i$ for single and multiple giants. Steps are median values, while vertical lines are 15.9–84.1% confidence intervals. The two distributions are not $\leq 2\sigma$ separated in all four considered ranges. 139

5.11 Cumulative $M \sin i$ distributions for multiple giants within 3 au and single giants split across an eccentricity of 0.4. While the single and circular giants are $> 4\sigma$ separated from the multis, eccentric giants are inconclusively more massive. Note that this may be a selection effect for more eccentric detections with greater mass. 140

5.12 Observed multiplicity of giant planets in the CLS sample. Half as many systems contain two or more giants as contain only one giant. . 141

5.13 $M \sin i$ distributions of giant planet pairs. We observe a 0.45 correlation in $\ln(M \sin i)$ between neighboring giants. 142

5.14 Observed semi-major axis ratio distribution of two-giant systems. The curve is a kernel density estimation of the distribution, while the vertical ticks show the observed set of axis ratios. Horizontal bar shows the kernel width. 144

LIST OF TABLES

Number		*Page*
2.1	Discovered or Revised Planets	34
3.1	Broken Power-Law Model Parameters	91
4.1	Small Planet Sample	106
4.2	Absolute and conditional probabilities and number of planets per star for inner small planets, outer giants, and Jupiter analogs.	111
4.3	An expanded version of Table 4.2's set of probabilities, with columns for small planets within 0.023–0.5 au and either 2–20 $M_\oplus$ or 3–30 $M_\oplus$.	120
5.1	Summary of population comparison between single giant systems and multi giant systems. Fraction of bound binaries comes from a literature estimate with Simbad and the Washington Double-star Catalog. Uncertainties come from Poisson counting error.	143

Chapter 1

INTRODUCTION

"God does not play dice with the universe; He plays an ineffable game of His own devising, which might be compared, from the perspective of any of the other players [i.e. everybody], to being involved in an obscure and complex variant of poker in a pitch-dark room, with blank cards, for infinite stakes, with a Dealer who won't tell you the rules, and who smiles all the time."

— Terry Pratchett, *Good Omens*

"I rejoice that I am to some extent restored to life by your work. If you had discovered any planets revolving around one of the fixed stars, there would now be waiting for me chains and a prison among Bruno's innumerabilities..."

— Johannes Kepler, *Kepler's Conversation with Galileo's Sidereal Messenger*

1.1 Is The Solar System 'Typical'?

What are the odds that the Solar System, Earth, and all life on it made it to this moment? How improbable is my existence? To answer those questions, we need to know how commonly habitable planets, defined as planets in orbits that can sustain liquid water (the 'Goldilocks Zone'), exist in the universe. That question is currently unanswerable within the limits of existing data, but we can build up to it by taking a census of many types of planets orbiting other stars, also known as extrasolar planets or exoplanets. This can bring us closer to understanding how typical or atypical the Solar System is. Are we one-in-a-million, or a common occurrence on cosmic scales?

In this introduction, I will set the stage for progress toward answering this question. I will review a history of humanity's understanding and discovery of planets, the current consensus of how planets are born and what is still not understood, and how we have and can still use statistical surveys of stars and planets to learn more about our place in the universe. This chapter draws from Hoskin, 2003 for a broad history of astronomy, from Seager and Lissauer, 2010 for a history of the discovery of exoplanets, and from Raymond and Morbidelli, 2022 for information on the state of planet formation theory.

1.2 The Nature of Planets

Our story begins at least five millennia before present day, when Egyptian priests tracked the motions of astronomical bodies to set the times of religious ceremonies and follow the gods in constellations (Hoskin, 2003). Babylonians within the next millennium developed new mathematics to keep time (leading to the sexagesimal clock system that we still use today), recorded the cycles of the Sun and Moon, and built models for their positions throughout the year. And Greek philosopher-scientists over a millennium after that used both empirical methods and conceptual reasoning to understand the world around them through astronomy, utilizing the Sun and trigonometry to measure the size of the Earth.

Throughout these eras, five points of light defied all possible explanations of their behavior: Mercury, Venus, Mars, Jupiter, and Saturn. These five 'stars' moved across unpredictable tracks through the sky, sometimes doubling back on themselves and then tracing out erratic loops. The Greeks called them "planets", named for a word meaning 'to wander around', because they wandered across the sky in a manner unlike other stars. The terracentric cosmology of Aristotle could not account for these irregular paths without arbitrarily complex models of concentric circles, orbits locked within orbits. Even then, reality as seen in the night sky diverged from predictions. This remained true through the Middle Ages, as Islamic scholars continued to refine astronomy for the sake of formalized times of prayer, designs for holy sites, and astrological predictions. The astronomers of that time and place inherited their systems from translations of Egyptian and Greek writings taken from the ruins of Alexandria and Constantinople. Although they developed ingenious mathematics and instruments to calculate the motions of astronomical bodies, their understanding of the five known planets remained bound by terracentrism and a belief that these points were just stars that happened to behave strangely.

The first real step toward understanding the nature of planets came from a Polish astronomer at the beginning of the 15th century. Nicholas Copernicus, a Rennaissance polymath who persisted in exploring the world beyond his obligations as a Catholic canon, recognized that the epicyclic models of Aristotle and Ptolemy lacked physical grounding and decided to invent a new paradigm. By 1514 he had published his

Little Commentary on a heliocentric theory of the solar system, which explained the changing brightnesses of the planets and the satellite nature of the Moon. In doing so, Copernicus placed the Earth as just one of a series of bodies orbiting the Sun and ordered the planets in distance and period.

Most European astronomers ignored Copernicus' claims and continued to twist older models into shapes that better fit existing observations, but the Dane Tycho Brahe and his apprentice Johannes Kepler performed new observations that led to more detailed descriptions of heliocentric orbits. Kepler in particular derived his eponymous laws of motion for the planets from their observations, finding that they (including the Earth) were bound to the Sun by comprehensible dynamics. Around the same time, Galileo Galilei struck two death blows to Aristotelian cosmology and to the consensus view of other planets. His innovative observational methods and brave defense of heliocentrism in the face of stubborn theocracy led to the eventual acceptance of heliocentrism, but perhaps just as important were his spatially resolved observations of Jupiter, Saturn, and Venus. His discovery of the moons of Jupiter and the rings of Saturn was proof that other planets are worlds unto themselves, with their own satellites, spheres of influence, and physical structure. His observations of the solar phases of Venus showed that Venus must be orbiting the Sun, and provided empirical disproof of Earth-centered cosmologies. Newton drove this point home within the next century by showing that Kepler's laws and the motions of the planets emerge from his proposed law of gravitational force. Astronomers understood the Solar System as a physically driven corpus, and the Earth lost its lonely status as a world.

1.3 Beyond Sol

As the Enlightenment swept over Europe, different thinkers had different reactions to the mounting evidence for Earth as one of many worlds. Even though Copernicus was the first known astronomer to propose and empirically investigate heliocentrism, he was reluctant to extrapolate the existence of planets orbiting other stars and chose to focus on what was provable with existing data and mathematics (Crowe, 1986). On the other end of the spectrum, the Italian friar Giordano Bruno published book after book in the 16th century advocating for the infinite nature of the universe and a plurality of worlds beyond Earth and the Solar System, and his theories only grew in influence after he was burned at the stake in Rome. Ironically, Kepler staunchly opposed the idea that there might be planets orbiting other stars, even while he revolutionized the orbital mechanics that astronomers still use today to characterize exoplanets. Immanuel Kant, one of the most important thinkers of the 18th century and any time before or after, argued in his *Universal Natural History and Theory of the Heavens* that an infinite universe governed by natural law inevitably would contain many worlds orbiting many stars (Kant, 1755). That same volume includes arguments in support of the Milky Way as the optical product of a dense disk of many stars, and a new theory of planet formation as the collapse of matter into increasingly dense due to gravity. Kant's push toward pluralism and a physics-based explanation of planet formation signals a post-Newton transition point toward naturally driven explanations of the world. This happened at the same time that astronomers began observing proof for these theories. William Herschel produced his stellar map of the Milky way in 1785, and his contemporaries searched for undiscovered stars and evidence of how they were born.

Eventually, astronomy expanded its focus from the search for and characterization of stars to the search for planets orbiting those stars. The quantitatively driven search for exoplanets first appeared in scientific literature in the mid-nineteenth century, when Jacob, 1855 claimed the existence of a giant planet orbiting a binary star system using a series of observations of the stellar positions. This also wound up becoming the first proven false-positive not long after, but the claim was notable as an indirect detection through the supposed planet's influence on the binary star's orbit. A few more astrometric false-positives came and went over the next century. Then, Struve, 1952 proposed a new method of detection: radial velocity (RV) observations of host stars. By using a ground-based spectrograph to measure the Doppler shift of a star over time, an observer could detect the periodic influence of a giant planet on the star's radial velocity and measure the planet's orbital period and minimum mass. Astronomers responded by reconfiguring existing spectrographs and builing new ones in order to conduct RV searches for giant planets. After a few brown dwarf detections and false-positives, Mayor and Queloz, 1995 detected the first known exoplanet orbiting a main-sequence star, a half-Jupiter-mass giant on a 4.2 day orbit. (Wolszczan and Frail, 1992 detected the first two known exoplanets using pulsar timing, but they orbit a dead star and are less relevant to the search.)

Over the next decade and a half, astronomers used new instruments and observing collaborations to detect more exoplanets via the radial velocity method, expanding to further-out orbits and smaller planets with more data and more precise measurements from new instruments. By the late 2000's, the astronomical community had discovered around 600 exoplanets, most of them via RV searches. The field began to branch out into the discovery of transiting exoplanets (also predicted by Struve, 1952) by monitoring the brightness of stars and searching for periodic dimming due to occultation. There were a few one-off detections starting with the near-simultaneous discoveries of a single transiting hot Jupiter (Charbonneau et al., 2000; Henry, Marcy, et al., 2000). Then the *Keper* space telescope (Borucki et al., 2010) began to survey a patch of 150,000 stars and blew the field wide open. The number of known exoplanets grew from several hundred to several thousand over the next decade, and surpassed 5,000 less than a month before the submission of this thesis. With the operation of the wide-field survey *TESS* space telescope (Ricker et al., 2014) over the last few years, as well as continuing and upcoming direct imaging and gravitational lensing surveys (Nielsen et al., 2019; Stassun et al., 2017), Earth-bound astronomers will continue to discover exoplanets for the foreseeable future.

1.4 Origin Story

A rapidly expanding and diversifying observed population of exoplanets naturally prompts theoretical investigation into the origins of these planets. As of 2022, astronomers' understanding of how planets are born is incomplete, but a few confidently vetted theories form a coherent story. This story begins with the formation of a star, within a cloud of gas and dust several light-years wide. As a molecular gas cloud collapses due to its own gravitational potential, it contracts more quickly along its axis of rotation, and flattens into a pancake-like structure with a newly-burning star at its center. This is known as a 'protoplanetary disk', because it is the birthplace of planets. This disk of gas and dust provides the material and environment needed to make planets.

First, dust grains in the disk a few microns wide coagulate into clumps. These clumps grow until they form small rocky bodies, ranging in size from pebbles a few centimeters across to larger masses a hundred kilometers across, known as planetesimals (Raymond and Morbidelli, 2022). Headwinds from the sub-Keplerian gas disk manifesting as Epstein drag slow solids down and push them together, leading to gravitational instabilities that collapse into kilometer-sized bodies (Youdin and Goodman, 2005; Hopkins and Christiansen, 2013). At the same time, these planetesimals accrete pebbles and grow to form larger rocky bodies, which hierarchically merge to form protoplanetary cores.

At this point the story branches out into a few paths. In one possible outcome, the core slows down in its growth as it consumes the rest of the material in its immediate environment. It completes its journey as a rocky planet with less than ten Earth masses of solids and a small fraction of that mass in a gas envelope. The core may accumulate a slightly higher fraction of gas, becoming a large planet with a high-metallicity core and gaseous outer envelope containing a few percent of the total mass budget. These planets are known as super-Earths or sub-Neptunes, depending on their composition.

More massive cores in the ten to fifteen Earth mass range can accrete larger amounts of gas. If the planet sits in a gas-rich environment and continues to grow until the mass of the envelope is comparable to the mass of the core, its gravitational potential exceeds the thermal energy of the local gas and runaway accretion begins. As the planet's gas envelope grows, it cools at a faster rate and contracts, which leads to more accretion, which leads to more contraction, and so on in a runaway process. The planet consumes all mass in some local regime until it becomes a gas giant, like Saturn or Jupiter. Our own solar system hosts planets spanning all of these possible outcomes, except for planets between the sizes of Earth and Neptune.

This story of planet formation leaves out a few important details, all of which are at least partly influenced by chance. Where in the disk are protoplanets most frequently born? Is it common for planets to migrate inward to tighter orbits as they grow, or are they more likely to stay put? What determines the final size of a planet, and what determines the distribution and mass ordering of planets (known as 'architecture') in a multi-planet system? These are some of the questions that I investigate in this thesis. The overarching question is, what does a representative sample of planetary systems look like, and why?

1.5 Counting to Answer Questions

In order to answer the above questions and discover insights into planet formation, we need to accurately and precisely infer an underlying distribution from an observed sample of planets. This requires a data-intensive survey to find planets, quantitative characterization of the sensitivity to discovering planets with those data, and careful statistics that combine these survey outputs to produce an expected planet distribution. Astronomers can test detailed theories of planet formation by comparing their predicted planet yields to the distribution that they actually measure. For instance, if Theory A predicts 10 Jupiter-like planets per 100 stars, and Theory B predicts 40 Jovians per 100 stars, one can use a Jovian-sensitive survey to measure the number of Jovians per star, otherwise known as the 'occurrence rate' of Jovians, and see whether it most closely matches Theory A, Theory B, or neither.

Work in the 90's such as Mayor and Queloz, 1995 and Marcy, Butler, Vogt, et al., 1998 discovered some of the first known exoplanets, but could not be used to measure occurrence rates because it only contained one-off discoveries not drawn from a broader survey. Starting around the turn of the millennium, several collaborations undertook unbiased radial velocity surveys of nearby (<100 pc) stars that are usable for planet occurrence measurements. These collaborations each selected a set of stars for radial velocity observations according to some objective criteria, such as brightness, stellar type, distance from the Solar System, and binary star status. They then aspired to observe of these stars with somewhat similar rates of measurements per time (cadence), regardless of whether or not they discovered evidence for planets in their data.

Naef et al., 2005 produced the first concrete occurrence measurements using a stellar sample in the double-digits. They observed a set of 330 stars with the ELODIE spectrograph (Baranne et al., 1996) to derive detection sensitivity contours in $M \sin i$ and orbital period space and infer the fraction of stars that host at least one giant planet. They found that $4 \pm 1\%$ of main-sequence stars host at least one giant planet more massive than half of Jupiter and within a 1500 day orbit. Fischer and Valenti, 2005 measured giant planet occurrence as a function of host star metallicity, and revealed that metal-richer stars are substantially more likely to host giant planets. This led to the insight that metal-richer stars are more likely to form large rocky cores that become giants, and is arguably the first time that occurrence measurements provided insights into the physics of planet formation. A few years later, Cumming et al., 2008 performed similar work to Naef et al., 2005, using the HIRES spectrograph to recover 48 planet candidates from a sample of 585 FGK stars and 8 years of HIRES data. They used this survey to make a more accurate measurement, and to report occurrence as a function of orbital period. They found an increase in occurrence beyond a one-year orbital period, near the water ice line for many main-sequence stars. This finding has led to well-supported theories that giant formation is influenced by some combination of solid accretion and x-ray sculpting of the early disk (Chachan, Lee, et al., 2021).

Since the late 00's, radial velocity occurrence work has continued at a steady pace, with more papers published as datasets and observational baselines grow. Mayor, Marmier, et al., 2011 and Howard, Marcy, Johnson, et al., 2010 provided independent measurements of the mass function of close-orbiting planets, showing that smaller planets are much more abundant than giants and contradicting the 'planet desert' predictions of consensus theories at the time (Ida and Lin, 2008a; Mordasini et al., 2009). Jones et al., 2016 and its continuation Wittenmyer, Wang, et al., 2020 used longer observational baselines to push confident measurements of occurrence as a function of orbital period out to 18 years, and confirmed the earlier finding that giant occurrence is much higher beyond 1 au than within.

An important aside: in practice these surveys were never as systematic and unbiased as they aspired to be. While the survey architects performed careful stellar selection, observational cadence for individual stars would almost always increase if the survey uncovered evidence for a planet orbiting that star. Even when this bias was mitigated, collaborations would still observe different stars with different cadences due to random factors, including nights available on a telescope, weather occluding certain parts of the sky, who constructed the starlist for a given night, and whether the observer on duty that night had a particular interest in a subset of stars. Sometimes these variables averaged out over time, but often they did not, and the results were surveys that were heterogeneous in both observations per star and time baseline. This heterogeneity in data turned into a heterogeneity in sensitivity to planets, which has become one of the principal challenges in using radial velocity surveys to infer occurrence rates from imperfectly conducted surveys.

The *Kepler* telescope circumvented the issue of survey variance and cleared a path to occurrence measurements for transiting planets. It observed a single patch of sky with uniform cadence for all 150,000 stars in its field of view for the first four years of its operation, producing a dataset with similar quality for all of its planned targets. The resulting planet catalog and uniform survey laid the foundation for a blitz of research into planet occurrence. Howard, Marcy, Bryson, et al., 2012 characterized the distribution of exoplanets with short orbital periods, this time with increased sensitivity to small planets and measurements of radii instead of $M \sin i$ as in prior RV work. Fulton, Petigura, Howard, et al., 2017 used careful measurements of stellar properties to more precisely measure close-in planet radii, and found a now-famous gap in the radius distribution around 1.5–2.0 $R_\oplus$, implying that there are two discrete types of rocky planets. Petigura, Marcy, and Howard, 2013, Foreman-Mackey, Hogg, and Morton, 2014, and Bryson et al., 2021 made independent attempts to measure the occurrence rate of Earth analogs orbiting Solar-like stars, trying to answer the question of how typical the Earth is and landing on an optimistic constraint of $0.88^{+1.28}_{-0.51}$ Earth analogs per G dwarf. Without *Kepler*, we would not have this first estimate of how unique our home is.

Beyond planet-by-planet statistics studies, much work has been done with *Kepler* on the nature of multi-planet system architectures. Fabrycky et al., 2014 reported on the orbital period distributions and ratios of multi-planet systems and found that planet pairs often linger just outside of resonant orbits, while Weiss et al., 2018 and Millholland and Winn, 2021 found that the sizes of planet pairs are correlated and suggest that disk properties strongly determine the final sizes of rocky cores. Work like this will continue to pull from the *Kepler* and *K2* samples, and the advent of *TESS* and its transiting planet yield will provide more opportunities to shed light on system architectures.

Although transit surveys have revolutionized our understanding of exoplanets, they are severely limited in sensitivity to long-period planets, because the probability of seeing a planet transit is inversely proportional to its distance from its host star. And although direct imaging and gravitational lensing surveys have improved our understanding of massive cold planets and small M-dwarf orbiters respectively (Nielsen et al., 2019; Biller et al., 2013; Clanton and Gaudi, 2016), imaging surveys are biased against less massive and 'older' ($\geq$ 1 Gyr) planets, while lensing surveys are severely limited by lack of knowledge of host star type, and have to use broad priors to infer stellar and therefore planetary parameters. Therefore, RV surveys are uniquely positioned to study the demographics of planets, particularly cold giants, orbiting stars that are not massive or young enough to be suitable for DI studies. It is with this focus in mind that astronomers have steadily incremented upon RV datasets over the past few decades, waiting to see new patterns emerge from growing planet catalogs. What can we learn about planetary system architectures with enough data, time, and patience?

1.6 Taking a New Census

After collecting over three decades of RV data with Lick-Hamilton, HIRES, and the Automated Planet Finder (APF), the California Planet Search team (CPS; Howard, Johnson, Marcy, Fischer, Wright, Bernat, et al., 2010) has laid the foundations for a substantially expanded census of exoplanets, one that pushes out to colder orbits than have have ever been characterized. As part of that team, I have had the privilege to make a new contribution to exoplanet statistics. In this thesis, I present a survey that is the culmination of several academic generations and over one hundred thousand

RV observations, and use that survey to drive insights into the formation pathways and dynamics of planets.

In Chapter 2, I present a high-precision RV survey of main-sequence stars, known hereafter as the California Legacy Survey (CLS). This work reports the stellar selection criteria for the CLS and the RV observations and associated data. I detail computational methods and open source software for the purpose of searching for planet candidates in and characterizing the sensitivity of RV data, and use these methods to produce a vetted catalog of 164 known and 14 newly discovered or revised exoplanets. We use the CLS catalog as a statistical sample for the work in all proceeding chapters.

In Chapter 3, BJ Fulton and I measure giant planet occurrence as a function of orbital separation. We confirm that giant planets are much more likely to exist beyond 1 au than within that distance, and find that our models favor a fall-off beyond ~8 au. This work also independently verifies occurrence measurements by direct imaging surveys of cold giants.

In Chapter 4, I investigate the relationship between cold giants and inner small planets and demonstrate that the conditional occurrences of both groups are significantly less than 100%. Colder and less massive giants are much more likely to host inner small planets than their warmer and more massive counterparts. I also confirm that the stellar hosts of small planets and giants are more metal-rich than the hosts of small planets without giant counterparts.

Finally, in Chapter 5, I take a comparative tour through the world of single and multiple giant planet systems. I characterize the orbital eccentricity distributions of these two groups and find that the lonely giants have a high-eccentricity tail not shared by the multiples. I also report that single giants have a super-hot Jupiter pile-up within 0.06 au not shared by multiples, that multiple giant hosts tend to be significantly metal-richer than lonely giant hosts, and that giants in the same system tend to have similar masses.

Chapter 2

TAKING A CENSUS: A CATALOG OF PLANETS DISCOVERED FROM RADIAL VELOCITY MONITORING

Rosenthal, L. J. et al. (July 2021). "The California Legacy Survey. I. A Catalog of 178 Planets from Precision Radial Velocity Monitoring of 719 Nearby Stars over Three Decades". In: *ApJS* 255.1, 8, p. 8. DOI: 10.3847/1538-4365/abe23c. arXiv: 2105.11583 [astro-ph.EP].

2.1 Introduction

Expanding and characterizing the population of known exoplanets with measured masses, orbital periods, and eccentricities is crucial to painting a more complete picture of planet formation and evolution. A census of diverse exoplanets sheds light on worlds radically different than Earth, and can provide insight into how these planets, as well our own Solar System, formed. For instance, the mass, semi-major axis, and eccentricity distributions of giant planets can be used to constrain formation scenarios for these objects. Nielsen et al., 2019 and Bowler, Blunt, et al., 2020 used mass and eccentricity constraints from direct imaging surveys to show that planetary-mass gas giants likely form via core accretion (Pollack et al., 1996), while more massive brown dwarfs and other substellar companions likely form via gravitational instability in protoplanetary disks (Boss, 1997). The present-day architectures and orbital properties of planetary systems can also be used to constrain their migration histories. Dawson and Murray-Clay, 2013 used a sample of giant planets with minimum masses and orbits constrained by radial velocity (RV) observations to provide evidence that giant planets orbiting metal-rich stars are more likely to be excited to high eccentricities or migrate inward due to planet-planet interactions. Many related questions remain unanswered. What is the mass–period distribution of planets out to 10 AU? How abundant are cold gas giants beyond the water ice line, and what can this abundance tell us about planet formation across protoplanetary disks? How do small, close-in planets arrive at their final masses and system architectures? What is the relationship between these warm small planets and cold gas giants; are their formation processes related? These questions can only be answered with an expansive and rigorously constructed census of exoplanets with measured masses and well-constrained orbits.

The community has made substantial progress on these fronts over the past few decades via targeted RV surveys. For instance, Bryan, Knutson, Howard, et al., 2016 surveyed 123 known giant hosts to study outer giant companions; they found that half of all giants have an outer companion, with tentatively declining frequency beyond 3 AU. Similarly, Knutson et al., 2014 found a 50% companion rate for transiting hot Jupiters using a sample of 51 stars. These two results suggest a planet formation process that favors giant multiplicity. On the small-planet front, Bryan, Knutson, Lee, et al., 2019 constructed an RV survey of 65 super-Earth hosts and found a giant companion rate of $39 \pm 7\%$. This suggests that these two populations are related in some way. Some questions have seen conflicting answers, requiring further work with a more expansive RV survey. For instance, Fernandes et al., 2019 studied planet occurrence as a function of orbital period by extracting the planetary minimum masses and periods, as well as completeness contours, from a catalog plot shown in Mayor, Marmier, et al., 2011, which presented a HARPS and CORALIE blind radial velocity survey of 822 stars and 155 planets over 10 yr (corresponding to a 4.6 AU circular orbit around a solar-mass star). The HARPS and CORALIE radial velocities were not published in Mayor, Marmier, et al., 2011, which measured giant planet occurrence as a function of orbital period out to 4000 days, in the range of the water ice line. Fernandes et al., 2019 pushed out to low-completeness regimes and estimated a sharp falloff in occurrence beyond the water ice line. In sharp contrast, Wittenmyer, Wang, et al., 2020 used their radial velocities from the Anglo-Australian Planet Search to construct a blind survey of 203 stars and 38 giant planets over 18 yr. They found that giant planet occurrence is roughly constant beyond the water ice line, out to almost 10 AU. The discrepancy between these two results needs to be resolved.

The California Planet Search team (CPS; Howard, Johnson, Marcy, Fischer, Wright, Bernat, et al., 2010) has conducted many RV surveys over the past three decades, in order to find exoplanets, measure their minimum masses, and characterize their orbits. Many of these surveys were designed explicitly for the purpose of studying planet occurrence. Therefore, they used stellar samples that were constructed without bias toward stars with known planets, or an increased likelihood of hosting planets, such as metal-rich stars (Gonzalez, 1997). For instance, the Keck Planet Search (Cumming et al., 2008) used 8 yr of Keck-HIRES data collected from 585 FGKM stars to study the occurrence of gas giants with periods as long as the survey baseline, measured the mass–period distribution of giant planets out to 5 AU, and found an increase in gas giant occurrence near the water ice line. The Eta-Earth

Survey (Howard, Marcy, Johnson, et al., 2010) used 5 yr of Keck-HIRES data collected from 166 Sun-like stars to measure the occurrence of planets with orbital periods less than 50 days, ranging from super-Earths to gas giants, and found both an abundance of planets within 10 day orbits and a mass function that increases with decreasing mass for close-in planets. The APF-50 Survey combined 5 yr of high-cadence Automated Planet Finder data on a sample of 50 bright, nearby stars with 20 yr of Keck-HIRES data to constrain the mass function of super-Earths and sub-Neptunes, and discovered several planets of both varieties (Fulton, Howard, et al., 2016).

We constructed an aggregate survey from these distinct RV surveys, known hereafter as the California Legacy Survey (CLS), in order to measure exoplanet occurrence, particularly for planets with long orbital periods. We selected every star in the CPS catalog that was observed as part of an occurrence survey, added 31 CPS stars that satisfied our stellar selection criteria (described below), and regularly observed these stars using the Keck and UCO-Lick observatories. The California Legacy Survey contains 103,991 RVs, and reaches observational baselines beyond three decades. We wrote an automated planet search pipeline to systematically recover all planets that are detectable in the CLS and to measure the search completeness of each star's RV time series. We can use these completeness contours to calculate exoplanet occurrence rates with respect to planetary and host-star properties (e.g. Cumming et al., 2008; Howard, Marcy, Johnson, et al., 2010).

In this chapter, we present the CLS stellar sample and the 164 known exoplanets orbiting these stars, as well as 14 newly discovered and vetted exoplanets and substellar companions. In Section 2, we describe our methodology for stellar selection. In Section 3, we describe the RVs measured for this survey. In Section 4, we describe our methods for computing the stellar properties of our sample. In Section 5, we describe the methods by which we search for exoplanets in the RVs, confirm their planetary status, and characterize their orbits. In Section 6, we present the catalog of known exoplanets, and describe in detail each of the new exoplanet candidates. In Section 7, we discuss the significance of our catalog, and conclude with plans for future work.

2.2 Stellar Sample Selection

Our goal for this study was to construct a sample of RV-observed FGKM stars and their associated planets, in order to provide a stellar and planetary catalog for occurrence studies. We want a survey that is quantifiably complete in some way, such as being volume- or magnitude-limited, so that we can perform unbiased occurrence measurements. One way to do this would be to observe every HD star within our desired range of stellar parameters, with the same cadence and a thirty year baseline. Given the constraints of finite observing time and instrumental magnitude limits, this is not possible. More importantly, there is no achievable, Platonic ideal of a quantifiably complete survey. However, we can approximate one by selecting CPS-observed stars that were originally chosen without bias toward a higher- or lower-than-average likelihood of hosting planets. Multiple CPS surveys, including the Keck Planet Search and Eta-Earth Survey, performed their stellar selection with these criteria.

We began with the Keck Planet Search sample, so that we can make direct comparisons to their results. We then supplemented those 585 stars with 135 stars that were not originally included as part of that sample, but they have since been observed by the CPS team and satisfy a set of criteria intended to ensure survey quality and statistical rigor for planet occurrence measurements. We selected these criteria to ensure data quality, both of individual measurements and stellar datasets, and proper stellar selection, without bias toward known or likely planet hosts, which would skew our occurrence measurements. We included CPS-observed stars that have at least 20 total RVs and at least 10 High Resolution Echelle Spectrometer (HIRES) RVs collected after the HIRES CCD upgrade in 2004, to guarantee enough RVs for well-constrained Keplerian fits, and have an observational baseline of at least 8 yr, which is the maximum baseline of the Cumming et al., 2008 sample at the time of publication. All stars in the Keck Planet Search sample pass these criteria, since we have collected more than 10 new HIRES RVs for each of them since 2004.

In order to ensure proper stellar selection, we did not include CPS-observed stars that were chosen for surveys that deliberately selected known planet hosts, metal-rich stars, or non-main-sequence stars, since these surveys would bias planet occurrence measurements. We excluded stars that were observed as part of the "N2K" and "M2K" surveys, which targeted metal rich stars to search for gas giants (Fischer, Laughlin, et al., 2005; Apps et al., 2010). We excluded all massive stars that were observed as part of a search for planets orbiting subgiants (Johnson, Bowler, et al., 2010), since that survey used a particular observing strategy geared solely toward detecting giant planets. We excluded all young stars that were selected for CPS observing based on photometric IR excess, since such stars were selected for an increased probability of planet occurrence (Hillenbrand et al., 2015). We excluded all stars from the "Friends of Hot Jupiters" surveys, which targeted known planet hosts (Knutson et al., 2014). For the same reason, we excluded all stars that were observed as part of *Kepler*, K2, TrES, HAT, WASP, or KELT transiting planet surveys (Bakos et al., 2002; Alonso et al., 2004; Pollacco et al., 2006; Pepper, 2007; Borucki, 2016).

This selection process left us with 719 stars. Figure 2.1 shows the entire CLS samples as a Venn diagram, illustrating the overlap of the Cumming et al., 2008 sample with the Eta-Earth (Howard, Marcy, Johnson, et al., 2010) and 25 pc northern hemisphere volume-limited (Hirsch et al., 2021) samples. The 25 pc sample includes 255 G and early K dwarfs with apparent V magnitudes ranging from $V \approx 3$ to $V \approx 9$. These stars have a median temperature of 5360 K and a median mass of 0.86 $M_{\odot}$. The median number and duration of RV observations for this sample was 71 RVs spanning 21 yr, while the minimum number and duration of observations in the sample was 20 RVs spanning 3 yr. The architects of all three of these surveys designed them for planet occurrence studies. Therefore, they did not construct these catalogs by selecting on properties known to correlate or anticorrelate with planet occurrence. There are only 31 stars in the California Legacy Survey that do not belong to any of these three surveys but do still pass of our selection criteria. This survey has no hard constraints on distance, apparent magnitude, or color, as seen in Figure 2.4.

2.3 Observations

Keck-HIRES

HIRES (Vogt, Allen, et al., 1994) has been in operation on the Keck I Telescope since 1994 and has been used to measure stellar RVs via the Doppler technique since 1996 (Cumming et al., 2008). This technique relies on measuring the Doppler shift of starlight relative to a reference spectrum of molecular iodine, which is at rest in the observatory frame (Butler, Marcy, Williams, et al., 1996). We consistently set up HIRES with the same wavelength format on the CCDs for each observation and followed other standard procedures of CPS Howard, Johnson, Marcy, Fischer, Wright, Bernat, et al., 2010. With the iodine technique, starlight passes through a glass cell of iodine gas heated to 50° C, imprinting thousands of molecular absorption lines onto the stellar spectrum, which act as a wavelength reference. We also collected an iodine-free "template" spectrum for each star. This spectrum is naturally convolved with the instrumental point spread function (PSF) and is

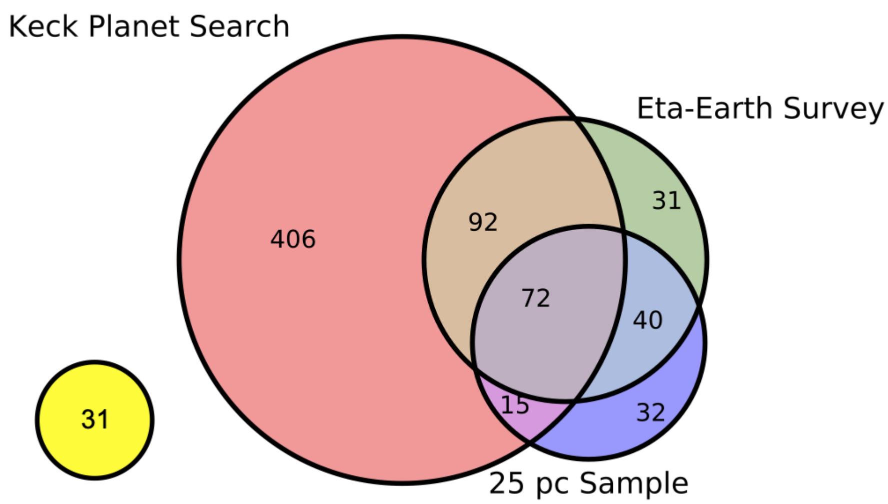

Figure 2.1: Venn diagram showing the overlap between the stars in the Keck Planet Search sample (Cumming et al., 2008), the Eta-Earth sample (Howard, Johnson, Marcy, Fischer, Wright, Bernat, et al., 2010), and a 25 pc northern hemisphere volume-limited survey (Hirsch et al. in prep). 31 stars in the California Legacy Survey do not belong to the union of these three surveys.

sampled at the resolving power of HIRES (R = 55,000–86,000, depending on the width of the decker used). These spectra are deconvolved using PSF measurements from spectra of featureless, rapidly rotating B stars with the iodine cell in the light path. The final, deconvolved intrinsic stellar spectra serve as ingredients in a forward-modeling procedure from which we measure relative Doppler shifts of each iodine-in spectrum of a given star (Valenti et al., 1995). We also used this process to compute uncertainties on the Doppler shifts. The uncertainty for each measurement is the standard error on the mean of the RVs for 700 segments of each spectrum (each 2 Å wide) run through the Doppler pipeline. We distinguish between "pre-upgrade" RVs (1996–2004; $\sim$3 m s^{-1} uncertainties) and "post-upgrade" RVs (2004–present; $\sim$1 m s^{-1} uncertainties). In 2004, HIRES was upgraded with a new CCD and other optical improvements. We account in the time series modeling for different RVs zero points (γ) for data from the two different eras.

The RVs reported here stem from HIRES observations with a long history. The RVs from 1996 to 2004 are based on HIRES spectra acquired by the California

& Carnegie Planet Search (CCPS) collaboration and were reported in Cumming et al., 2008. CCPS continued to observe these stars, but split into two separate collaborations: CPS and the Lick-Carnegie Exoplanet Survey (LCES). This work principally reports results from 41,804 CPS and CCPS HIRES spectra that were obtained and analyzed by our team during 1996–2020. In addition, we have included RVs computed by our pipeline for 7530 spectra of CLS stars taken by LCES during 2008–2014. These HIRES spectra were acquired with the same instrumental setup as the CPS spectra and are publicly available in the Keck Observatory Archive. Butler, Vogt, et al., 2017 separately published RVs based on the same HIRES observations from CCPS, CPS, and LCES for the 1996–2014 time span. The LCES and CPS Doppler pipelines diverged in ~2007. Tal-Or et al., 2019 uncovered the 2004 zero-point offset, which we model with two independent offsets. They also claimed two second-order systematics in the LCES 2017 dataset: a long-term drift of order $< 1\ \mathrm{m\,s^{-1}}$, and a correlation between stellar RVs and time of night with respect to midnight. They estimated the long-term drift by averaging the zero points of three RV-quiet stars on each night, where possible. However, by our estimates, even the quietest stars exhibit 1–2 $\mathrm{m\,s^{-1}}$ jitter in HIRES time series. Averaging the zero points of three such stars will likely yield a scatter of $1\ \mathrm{m\,s^{-1}}$ across many nights. Additionally, they did not remove planet RV signals from their data before estimating the linear correlation between RV and time of night, and it is unclear how they derived the uncertainty in that correlation.

Automated Planet Finder

The APF-Levy spectrograph is a robotic telescope near the summit of Mt. Hamilton, designed to find and characterize exoplanets with high-cadence Doppler spectroscopy (Vogt, Radovan, et al., 2014; Radovan et al., 2014). The facility consists of a 2.4-m telescope and the Levy Spectrometer, which has been optimized for optical Doppler shift measurements. The Doppler pipeline that was developed for Keck-HIRES also extracts RV measurements from APF spectra. Most of the APF data in the California Legacy Survey was collected as part of the APF-50 Survey (Fulton, Petigura, Howard, et al., 2017), the stellar sample of which was drawn entirely from the Eta Earth sample. These two surveys have slightly different selection criteria. While both surveys have a distance cut $d < 25$ pc and luminosity cut $M_V < 3$, Eta-Earth cuts on apparent magnitude $V < 11$, whereas APF-50 has $V < 7$; Eta-Earth cuts on chromospheric activity $\log R'_{\mathrm{HK}} < -4.7$, whereas APF-50 has $\log R'_{\mathrm{HK}} < -4.95$; and Eta-Earth cuts on declination $> -30°$, whereas APF-50 has declination $> -10^\circ$. These stricter cuts were made to ensure higher data quality for the high-cadence APF survey.

Lick-Hamilton

The Hamilton Spectrograph is a high-resolution echelle spectrometer, attached to the 3 m Shane telescope on Mt. Hamilton. Beginning in 1987, and ending in 2011 with a catastrophic iodine cell failure, the Lick Planet Search program (Fischer, Marcy, et al., 2014) monitored 387 bright FGKM dwarfs to search for and characterize giant exoplanets. This was one of the first surveys to produce precise RVs via Doppler spectroscopy with iodine cell calibration, and yielded RVs with precision in the range 3–10 $\mathrm{m\,s^{-1}}$. The Lick Planet Search overlaps heavily with the Keck Planet Search and other CPS surveys, since these surveys drew from the same bright-star catalogs.

Activity Indices

For each HIRES and APF spectrum from which we measure radial velocities, we also measure the strength of emission in the cores of the Ca II H & K lines (S-values) following the techniques of Isaacson and Fischer, 2010 and Robertson, Mahadevan, et al., 2014. There is a small, arbitrary offset between the HIRES and APF activity indices. We adopted uniform S-value uncertainties with values of 0.002 and 0.004 for HIRES and APF respectively. We provide activity indices along with our RV measurements. Missing values are the result of sky contamination and/or low SNR.

APT Photometry

We collected long-term photometric observations of the subset of our sample that were included in the APF-50 survey (Fulton, Petigura, Howard, et al., 2017), in order to search for evidence of rotation-induced stellar activity. We collected these measurements with Tennessee State University's Automated Photometric Telescopes (APTs) at Fairborn Observatory as part of a long-term program to study stellar magnetic activity cycles (Lockwood et al., 2013). Most stars have photometric datasets spanning 15 – 23 yr. The APTs are equipped with photomultiplier tubes that measure the flux in the Stromgren b and y bands relative to three comparison stars. We combined the differential b and y measurements into a single $(b + y)/2$ "passband" then converted the differential magnitudes into a relative flux normalized to 1.0. The precision in relative flux is typically between 0.001 and 0.0015. Further details of the observing strategy and data reduction pipeline are available in Henry, 1999; Eaton et al., 2003; Henry, Kane, et al., 2013. We make the photometric data available as a machine-readable table at github.com/leerosenthalj/CLSI.

Observational Statistics

We examined the range of observing cadences and observational baselines within the CLS sample, to determine whether stars without known planets were observed with strategies that differed significantly from those for stars with known planets. Figure 2 shows the distribution of number of observations and observational baselines for three groups of stars: the entire sample, the stars around which we detected planets, and the star around which we did not detect planets. Each of these three samples has a median baseline of 21 yr. Stars with detected planet have a median of 74 observations, compared to 35 observations for stars without detected planets and 41 observations for the entire CLS sample. A factor of two in number of observations will have a small but measurable impact on planet detectability of a given data set, and therefore on its search completeness contours.

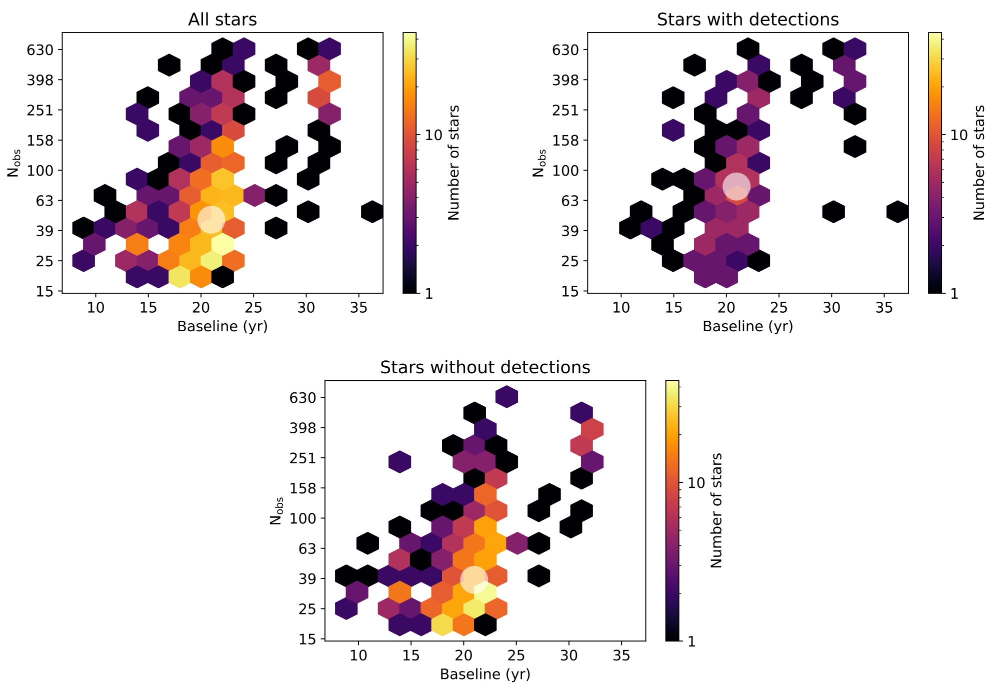

Figure 2.2: Distributions of observational baseline versus number of observations. The top left panel shows these statistics for all stars in the CLS sample; the top right panel shows stars around which we detect planets; the bottom panel shows stars around which we do not detect planets. Median baseline and number of observations for each sample are overplotted as translucent circles.

2.4 Stellar properties

We derived stellar properties for our sample by applying the `SpecMatch` (Petigura, 2015) and `Isoclassify` (Huber, 2017) software packages to the template Keck-HIRES spectra of our stars. `Specmatch` takes an optical stellar spectrum as input, and by interpolating over a grid of template spectra with known associated stellar properties, returns three spectral properties and uncertainties. For stars hotter than 4700 K, we interpolated over synthetic spectra to derive spectral parameters (Petigura, 2015). For stars below this threshold, we interpolated over real spectra of cool stars with well-characterized stellar properties, since synthetic spectral models are unreliable below this temperature (Yee et al., 2017).

`Specmatch` produces metallicity, effective temperature, and surface gravity when interpolating over synthetic spectra; it produces metallicity, effective temperature, and radius when interpolating over empirical spectra. `Isoclassify` takes effective temperature, metallicity, and surface gravity as spectral parameter inputs, and uses isochrone models and multinest Bayesian sampling (Buchner, 2016) to produce estimates and uncertainties of physical parameters, in particular stellar mass. For stars cooler than 4700 K, we passed `Isoclassify` a wide Gaussian input prior on surface gravity, since temperature and metallicity strongly constrain the masses of cool, main-sequence stars (Johnson, Petigura, et al., 2017).

Almost all stars in the California Legacy Survey have both Gaia-measured parallaxes (Gaia Collaboration, Brown, Vallenari, Prusti, de Bruijne, Mignard, et al., 2016; Gaia Collaboration, Brown, Vallenari, Prusti, de Bruijne, Babusiaux, et al., 2018; Lindegren et al., 2018) and apparent K-band magnitudes. For stars with both of these measurements available, we pass them and their uncertainties into `Isoclassify` as additional inputs, since taken together, they constrain stellar luminosity and therefore place tighter constraints on stellar mass. `Isoclassify` also returns more precise estimates of stellar radius when provided with parallax and apparent magnitude. With the inclusion of this luminosity constraint, the median precision of our stellar mass measurements is 3.6%.

In Table 2 in Appendix B, we report stellar mass, radius, surface gravity, effective temperature, and metallicity for a subsection of the CLS sample. We make this table available for the entire sample at github.com/leerosenthalj/CLSI, with additional columns including V-band magnitude and Gaia parallax. Figure 3 is a visualization of these stellar properties, while Figure 4 shows individual histograms for mass, metallicity, and effective temperature, as well as for the following observational properties: parallax-inverse distance, V, and $B - V$.

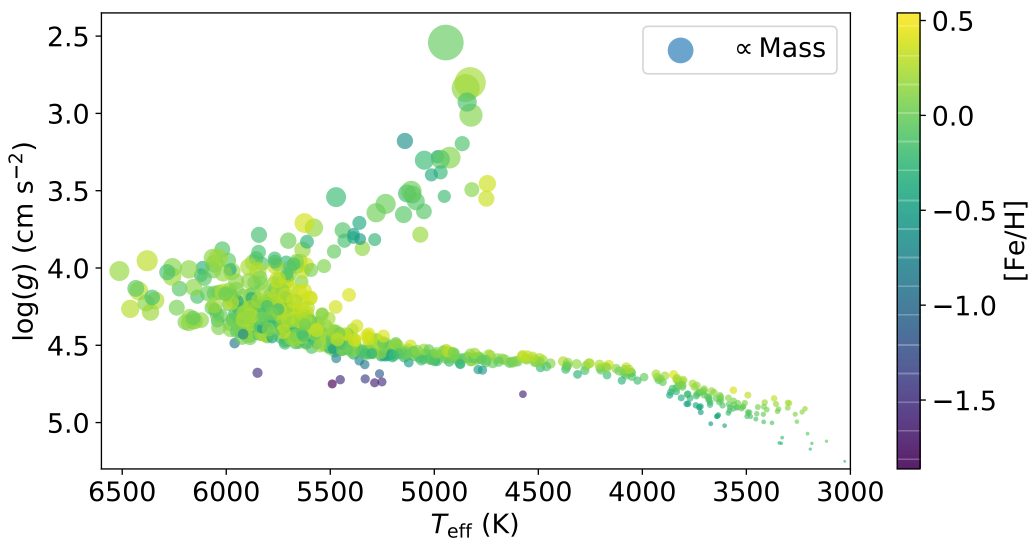

Figure 2.3: Stellar property measurements of the California Legacy Survey, in effective temperature, surface gravity, and mass. The sample consists of stars spanning spectral types F, G, K, and M, some of which have evolved off of the main sequence. Most stars have metallicities within 0.4 dex of Solar metallicity, with the exception of a small handful of extremely metal-poor stars, which lie below the main sequence on this plot.

2.5 Planet Catalog Methods

Planet Search

We developed an iterative approach to a search for periodic signals in RV data in order to generate the CLS planet catalog. We outline this algorithm, which we developed as the open-source Python package `RVSearch` and have made public alongside the publication of Rosenthal, Fulton, et al., 2021. Figure 2.5 is a flowchart that lays out each step of the algorithm, and Figure 2.6 is a visualization of an example `RVSearch` output, where the top two panels show the final model, and each successive row shows an iterative search for each signal in the model. First, we provide an initial model, from which the iterative search begins. This initial model contains an RV data set, and a likelihood function. The natural logarithm of the latter is defined as

$$\ln(\mathcal{L}) = -\frac{1}{2}\sum_i \left[\frac{(v_i - m(t_i) - \gamma_D)^2}{\sigma_i^2} + \ln(2\pi\sigma_i^2)\right], \tag{2.1}$$

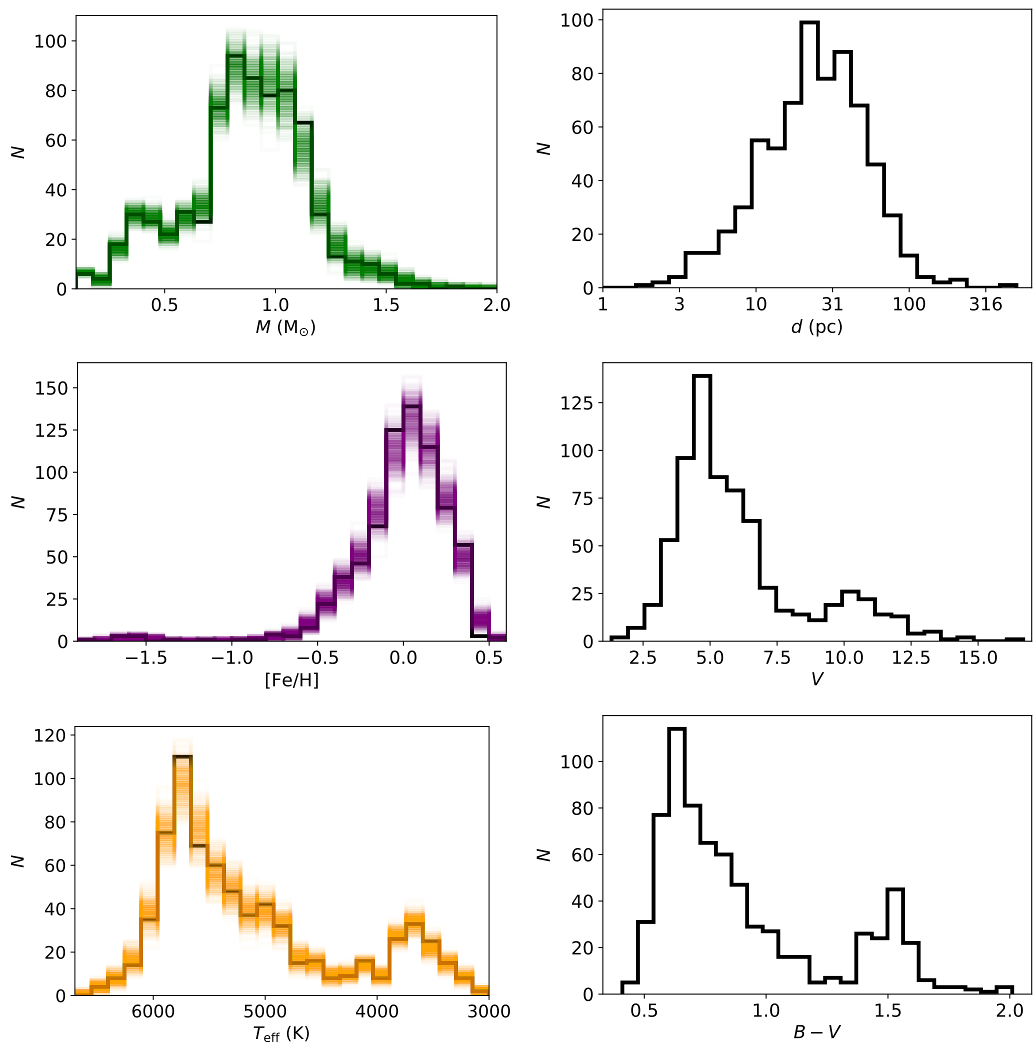

Figure 2.4: Stellar parameter distributions. The left column shows mass, metallicity, and effective temperature, while the right column shows parallax-inferred distance, V-band magnitude, and $B-V$ color. Black lines are histograms of the stellar parameter median values. For the left column, colored lines are 500 histograms per panel, with parameters redrawn from normal distributions with width equal to their individual measurement uncertainties. We left these redrawn parameter histograms out for the plots in the right column because distance, magnitude, and color have uncertainties that are smaller than the chosen bin size.

where i is the measurement index, v_i is the ith RV measurement, γ_D is the offset of the instrumental dataset from which the ith measurement is drawn, and σ_i^2 is the quadrature sum of the instrumental error and the stellar jitter term of the ith measurement's instrumental data set. Here, $m(t_i)$ is the model RV at time t_i, defined as

$$m(t) = \sum_n \mathrm{K}(t|K_n, P_n, e_n, \omega_n, t_{cn}) + \dot{\gamma}(t - t_0) + \ddot{\gamma}(t - t_0)^2, \tag{2.2}$$

where n is a given Keplerian orbit in the model, $\mathrm{K}(t|K, P, e, \omega, t_c)$ is the Keplerian orbit RV signature at time t given RV amplitude K, period P, eccentricity e, argument of periastron ω, and time of inferior conjunction t_c, $\dot{\gamma}$ is a linear trend term, $\ddot{\gamma}$ is a quadratic trend term, and t_0 is a reference time, which we defined as the median time of observation.

We used `RadVel` (Fulton, Petigura, Blunt, et al., 2018) to fit Keplerian orbits. The initial likelihood model contains either a one-planet Keplerian model with undefined orbital parameters, or a predefined model including trend/curvature terms and/or Keplerian terms associated with known orbital companions. We defaulted to performing a blind search starting with the undefined single-planet model, and we only supply a predefined model if there is evidence for a highly eccentric companion whose period is misidentified by our search algorithm. Several highly eccentric stellar binaries satisfy this criterion, as do two planets: HD 120066 b (Blunt et al., 2019), and HD 80606 b (Wittenmyer, Endl, Cochran, and Levison, 2007).

Before beginning a blind search, `RVSearch` determines whether the data merits a trend with curvature, a linear trend, or no trend. It does this by fitting each of these three models to the data, then performing a goodness-of-fit test to decide which model is favored. We measured the Bayesian Information Criterion (BIC) for each of the three models, and computed the ΔBIC between each model. RVSearch selects the linear model if it has $\Delta\mathrm{BIC} = 5$ with respect to the flat model, and the quadratic model if it has $\Delta\mathrm{BIC} = 5$ with respect to the linear model. We did not perform this test on datasets that contain eccentric companions with orbital periods greater than the data's observational baseline, since such datasets would be better fit with a long-period Keplerian orbit than with linear and parabolic trends. The Bayesian information criterion is defined as

$$\mathrm{BIC} = k\ln(n_{\mathrm{obs}}) - 2\ln(\mathcal{L}), \tag{2.3}$$

where n_{obs} is the number of observations, k is the number of free model parameters, and $\ln(\mathcal{L})$ is the log-likelihood of the model in question.

Once we provide an initial model, `RVSearch` defines an orbital period grid over which to search, with sampling such that the difference in frequency between adjacent grid points is $\frac{1}{2\pi\tau}$, where τ is the observational baseline. We chose this grid spacing in accordance with Horne and Baliunas, 1986, who state that, in frequency space, a Lomb–Scargle periodogram has a minimum peak width of $\frac{1}{2\pi\tau}$. For each dataset, we searched for periodicity between two days and five times the observational baseline. Searching out to five times the baseline only adds a few more points to the period grid, and it allows for the possibility of recovering highly eccentric, ultra-long-period planet candidates with best-fit orbital period.

The search algorithm then computes a goodness-of-fit periodogram by iterating through the period grid and fitting a sinusoid with a fixed period to the data. We measure goodness-of-fit as the ΔBIC at each grid point between the best-fit, n+1-planet model with the given fixed period, and the n-planet fit to the data (this is the zero-planet model for the first planet search).

After constructing a ΔBIC periodogram, the algorithm performs a linear fit to a log-scale histogram of the periodogram power values. The algorithm then extrapolates a ΔBIC detection threshold corresponding to an empirical false-alarm probability of 0.1%, meaning that, according to the power-law fit, only 0.1% of periodogram values are expected to fall beyond this threshold. This process follows the detection methodology outlined in Howard and Fulton, 2016.

If a periodic signal exceeds this detection threshold, `RVSearch` refines the fit of the corresponding Keplerian orbit by performing a maximum *a posteriori* (MAP) fit with all model parameters free, including eccentricity, and records the BIC of that best-fit model. `RVSearch` includes two hard-bound priors, which constrain $K > 0$ and $0 <= e < 1$. The algorithm then adds another planet to the RV model and conducts another grid search, leaving all parameters of the known Keplerian orbits free so that they might converge to a more optimal solution. In the case of the search for the second planet in a system, the goodness of fit is defined as the difference between the BIC of the best-fit one-planet model and the BIC of the two-planet model at each fixed period in the grid. `RVSearch` once again sets a detection threshold in the manner described above, and this iterative search continues until it returns a nondetection.

This iterative periodogram search is superior to a Lomb–Scargle residual subtraction search in two key ways. First, this process fits for the instrument-specific parameters of each dataset, stellar jitter and RV-offset, as free parameters throughout the search. Second, by leaving the known model parameters free while searching for each successive planet, we allow the solutions for the already discovered planets to reach better max-likelihood solutions that only become evident with the inclusion of another planet in the model.

Note that our search and model comparison process is not Bayesian; we do not use priors to inform our model selection, and we do not sample posteriors, beyond a grid search in period space, until we settle upon a final model. We use the BIC as our model comparison metric because it incorporates the number of free parameters as a penalty on more complex models, which, in our case, corresponds to models with additional planets.

We make `RVSearch` publicly available via a GitHub repository. See the RVSearch website for installation and use instructions.

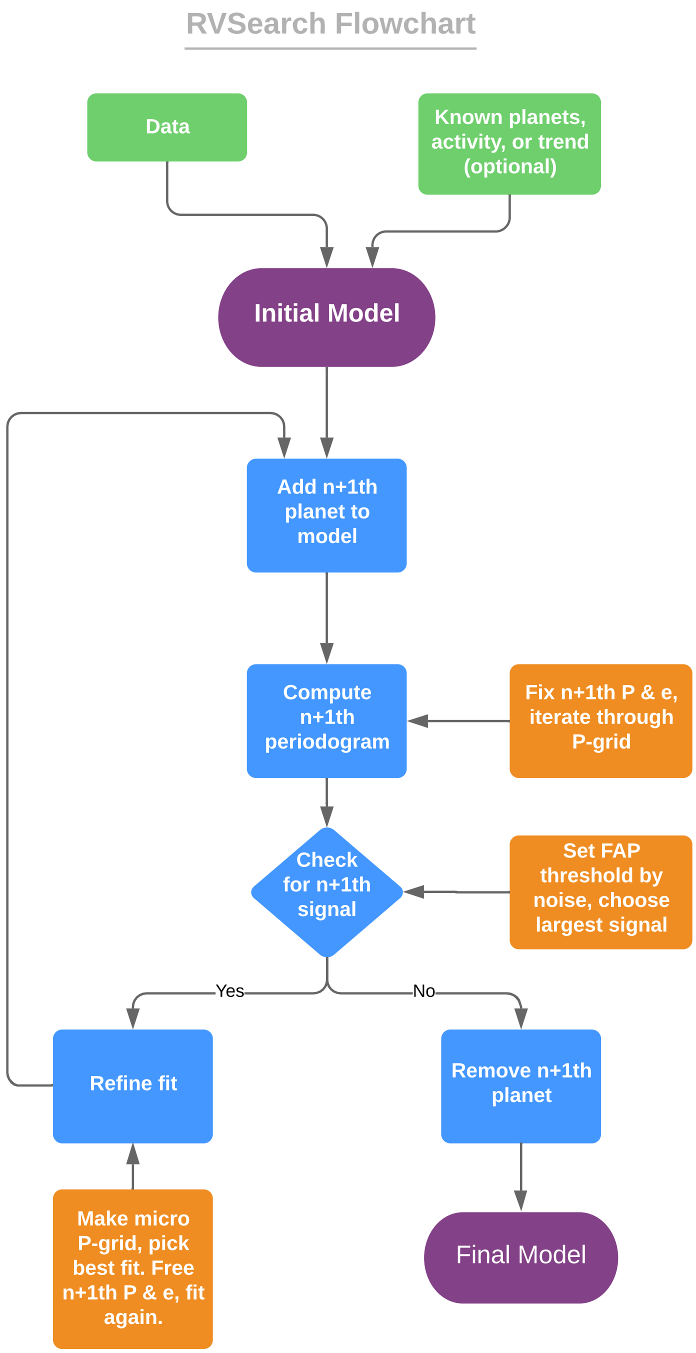

Figure 2.5: Search algorithm flowchart.

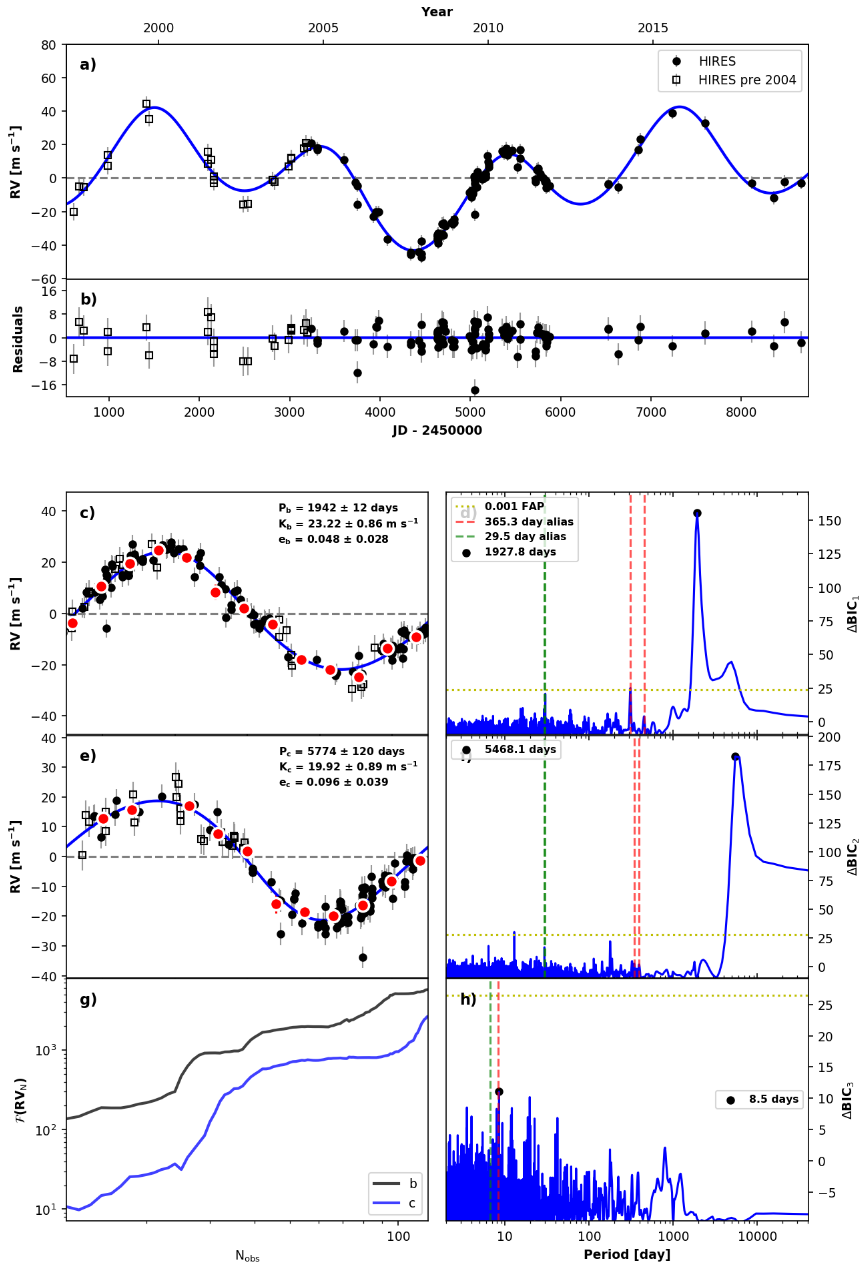

Figure 2.6: Example RVSearch summary plot, for the known two-planet system HIP 109388. Panel a) shows the total model plotted over the radial velocity time series, while panel b) shows the model residuals. Each successive row shows a phase-folded signal discovered by RVSearch on the left, and the associated periodogram on the right. The final row shows the running periodograms of each signal, generated with Lomb–Scargle power, on the left, and the final periodogram on the right.

Search Completeness

We characterized the search completeness of each individual dataset, and of the entire survey, by running injection-recovery tests. Once `RVSearch` completed an iterative search of a dataset, it injected synthetic planets into the data and ran one more search iteration to determine whether it recovers these synthetic planets in that particular dataset. We ran 3000 injection tests for each star. We drew the injected planet period and $M \sin i$ from log-uniform distributions, and drew eccentricity from the beta distribution described in Kipping, 2013, which was fit to a population of RV-observed planets.

We used the results of these injection tests to compute search completeness for each individual dataset, and report constant $M \sin i$ / a contours of detection probability. Figure 2.7 shows examples of these contours and the corresponding RVs for three different stars, all early G-type: one with 25 observations, one with 94, and one with 372. We make the 10th, 50th, and 90th percentile completeness contours for each individual star available at `github.com/leerosenthalj/CLSI`.

Model posteriors

Once `RVSearch` returned max-likelihood estimates of the orbital model parameters for a given dataset, we sampled the model posterior using affine-invariant sampling, implemented via `emcee` and `RadVel` (Foreman-Mackey, Hogg, Lang, et al., 2013b; Fulton, Petigura, Blunt, et al., 2018). We sampled using the orbital parameter basis $\{ \log P \; K \; t_c \; \sqrt{e}\sin\omega \; \sqrt{e}\cos\omega \}$. We placed uniform priors on all fitting parameters, with hard bounds such that $K > 0$ and $0 \leq e < 1$. We fit in $\log P$ space to efficiently sample orbits with periods longer than our observational baseline, and in $\sqrt{e}\sin\omega$ and $\sqrt{e}\cos\omega$ to minimize bias toward higher eccentricities (Lucy and Sweeney, 1971). We reported parameter estimates and uncertainties as the median and $\pm 1\sigma$ intervals.

If a dataset is so poorly constrained by a Keplerian model that `emcee`'s affine-invariant sampler cannot efficiently sample the posterior distribution, we instead used a rejection sampling algorithm to estimate the posterior. In these cases, we used `TheJoker` (Price-Whelan et al., 2017), a modified MCMC algorithm designed to sample Keplerian orbital fits to sparse radial velocity measurements. We chose a flat prior on $\log P$, with a minimum at the observing baseline and a maximum at twenty times the observing baseline. We drew orbital eccentricity from a beta prior weighted toward zero, as modeled in Kipping, 2013, in order to downweight orbits

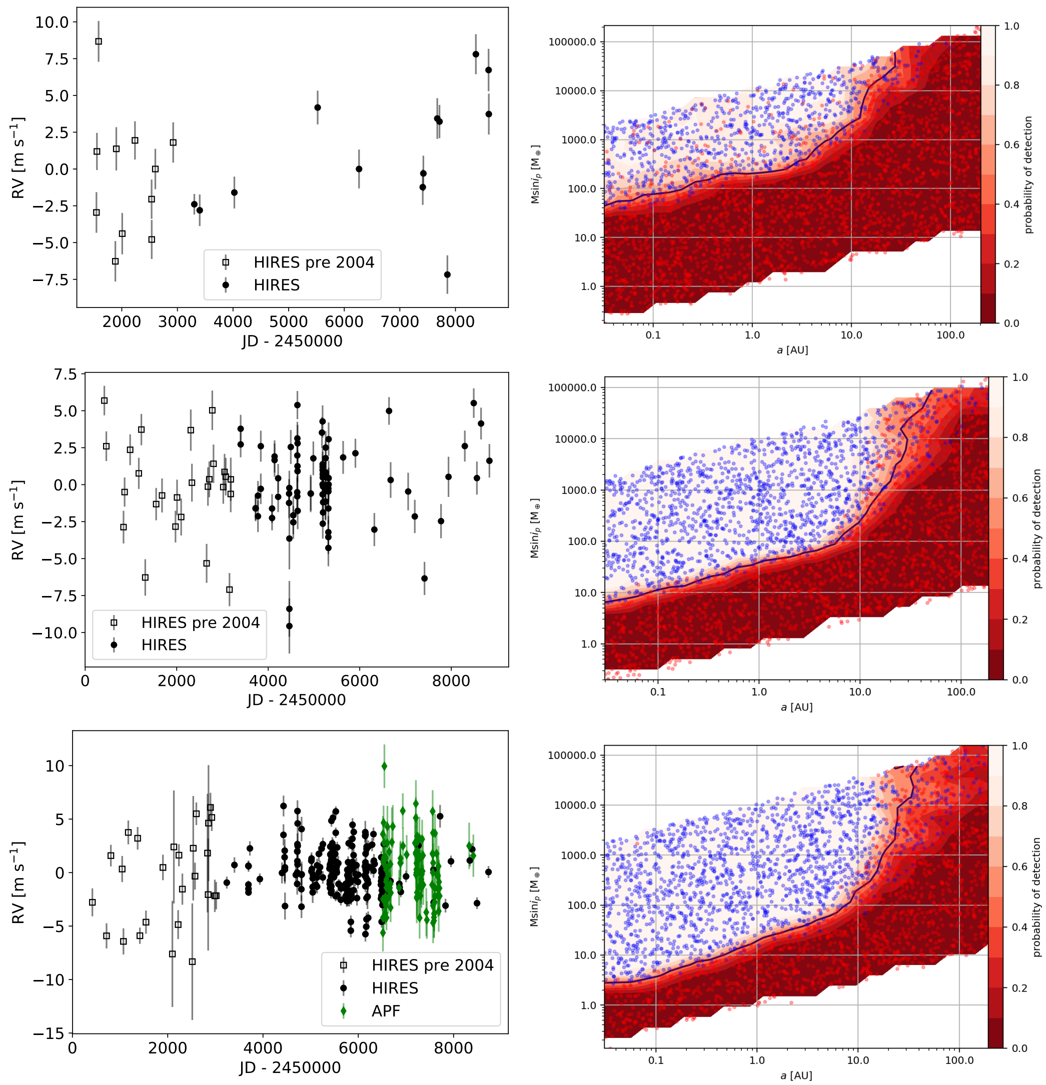

Figure 2.7: RVs and completeness contours for three datasets with similar baselines, median measurement errors, and stellar jitter. The left column plots RVs with respect to time, while the right column plots injected signals in the $M \sin i$ and a plane, where blue dots are recovered injections and red dots are not. The right column also shows detection probability contours, with 50% plotted as a solid black line. From top to bottom, we show RVs and contours for HD 44420, for which we have 24 RVs; HD 97343, for which we have 94 RVs; and HD 12051, for which we have 372 RVs.

with arbitrarily high eccentricity, which can be viable fits to sparse or otherwise underconstraining RV data sets.

False-positive vetting

We performed a series of tests to vet each planet candidate discovered by our search pipeline. The following subsections each detail one test we perform to rule out one way in which a signal might be a false-positive. We also represent this process with a flowchart in Figure 2.9, and include a table of all false-positive signals recovered by `RVSearch` in Table 6.

Stellar activity, magnetic/long-period

Many main-sequence stars, particularly F- and G-type, have magnetic activity cycles on timescales of several to tens of years. These fluctuations in activity can cause changes in the core depths of stellar Calcium H & K lines, which manifest as apparent RV shifts (Isaacson and Fischer, 2010). To evaluate whether stellar activity may be the cause of a signal recovered by our search pipeline, we measure the linear correlation between the RV signature of that signal and a measured stellar activity metric–in our case, S-values. We computed S-values for both post-upgrade HIRES and APF data by measuring the core flux of Calcium H & K lines.

If we found a periodic signal in the S-value data that has a similar period and phase similar to one of the Keplerian terms in our RV model, we searched for correlations between our RV model and S-values. If we found one periodic signal in an RV dataset, we measured its correlation with stellar activity simply as the linear correlations between the RVs of each instrument and their associated S-values. If we found multiple periodic signals, then for each signal, we subtracted the associated RV models of all other signals from the data, and measured the correlations between these residuals and the S-values. A significant linear correlation between a signal's RV residuals and the associated S-values does not necessarily mean that this signal is caused by stellar activity, even when these signals also have the same period and phase, but we took it as sufficient evidence to remove such signals from our catalog of confirmed planets.

It is important to note that our approach to vetting our planet candidates is systematic but not exhaustive, particularly with respect to stellar activity. One might use activity metrics beyond S-values and photometry, such as Hα line modulation. Furthermore, there are more sophisticated ways to deal with activity than searching for linear correlations with RVs. For instance, one might actively model stellar activity during the search process, using a Gaussian Process (Haywood et al., 2014) or some other correlated noise model. Such techniques might improve the accuracy of our planet candidate parameters and catalog selection, but require case-by-case analysis for each stellar system, as activity modeling is sometimes unwarranted or even counterproductive, e.g., for low-activity stars or confirmed planets that have periods similar to their host star's activity cycle. We chose to perform uniform, after-the-fact vetting for our catalog, and invite others to perform more sophisticated modeling for individual systems of interest.

Stellar activity, rotation/short-period

We only detected planet candidates that are low-amplitude and short-period enough to possibly be stellar rotation false positives in sustained, high-cadence datasets. Almost all CLS datasets that satisfy this criteria were collected as part of the APF-50 survey. We collected APT photometry of all APF-50 stars, which we can use to search for evidence of stellar rotation with moving-average smoothing and periodogram analysis. If we find strong evidence for rotation in APT photometry, or spectral S-value measurements, we discount planet candidates with periods close to the apparent rotation timescale or its harmonics.

Yearly alias

When we find a signal with a period of a year or an integer fraction of a year, we investigate whether it is an alias of long-period power, or a systematic that is correlated with the barycentric velocity at the time of observation or Doppler fitting parameters. We do this by recomputing the associated RVs using a different template observation. When another template observation was unavailable, we were able to take one using Keck-HIRES during collaborator observing nights. Templates taken in poor observing conditions or when barycentric velocity with respect to the observed star is high can produce systematic errors in the Doppler code. If a search of this new dataset returns a nondetection, or detection at a significantly different period, we conclude that this signal is an alias. Figure 8 shows the presence of yearly alias power in our survey, seen in a stack of the the final nondetection periodograms of all CLS stars.

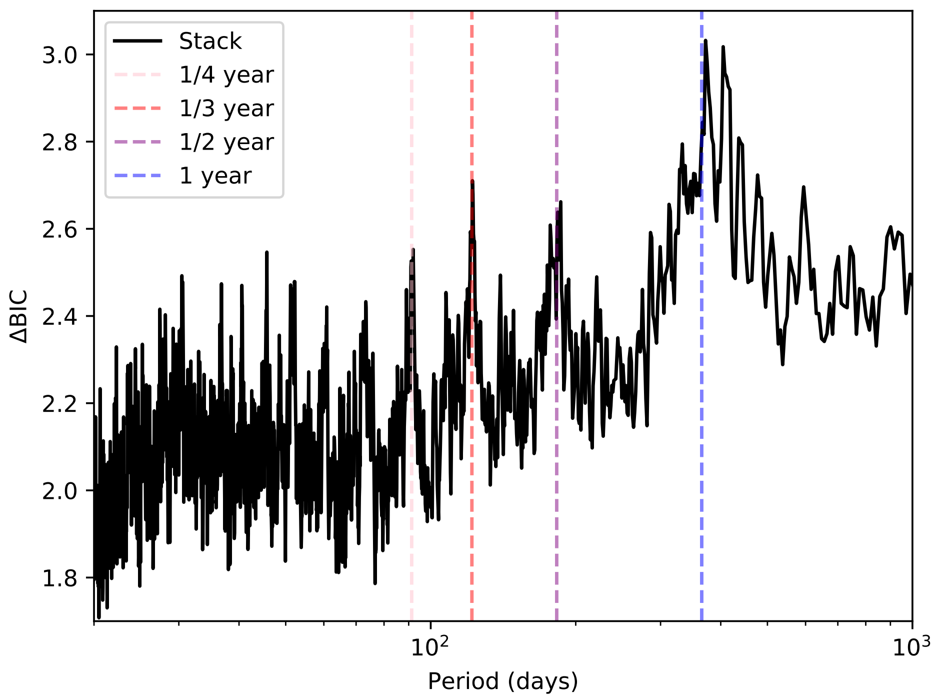

Figure 2.8: Stack of all final nondetection periodograms in the CLS planet search, linearly interpolated to the same period grid. A broad peak around 1 year is evident, as well as narrow peaks at 1/2-year, 1/3-year, and 1/4-year.

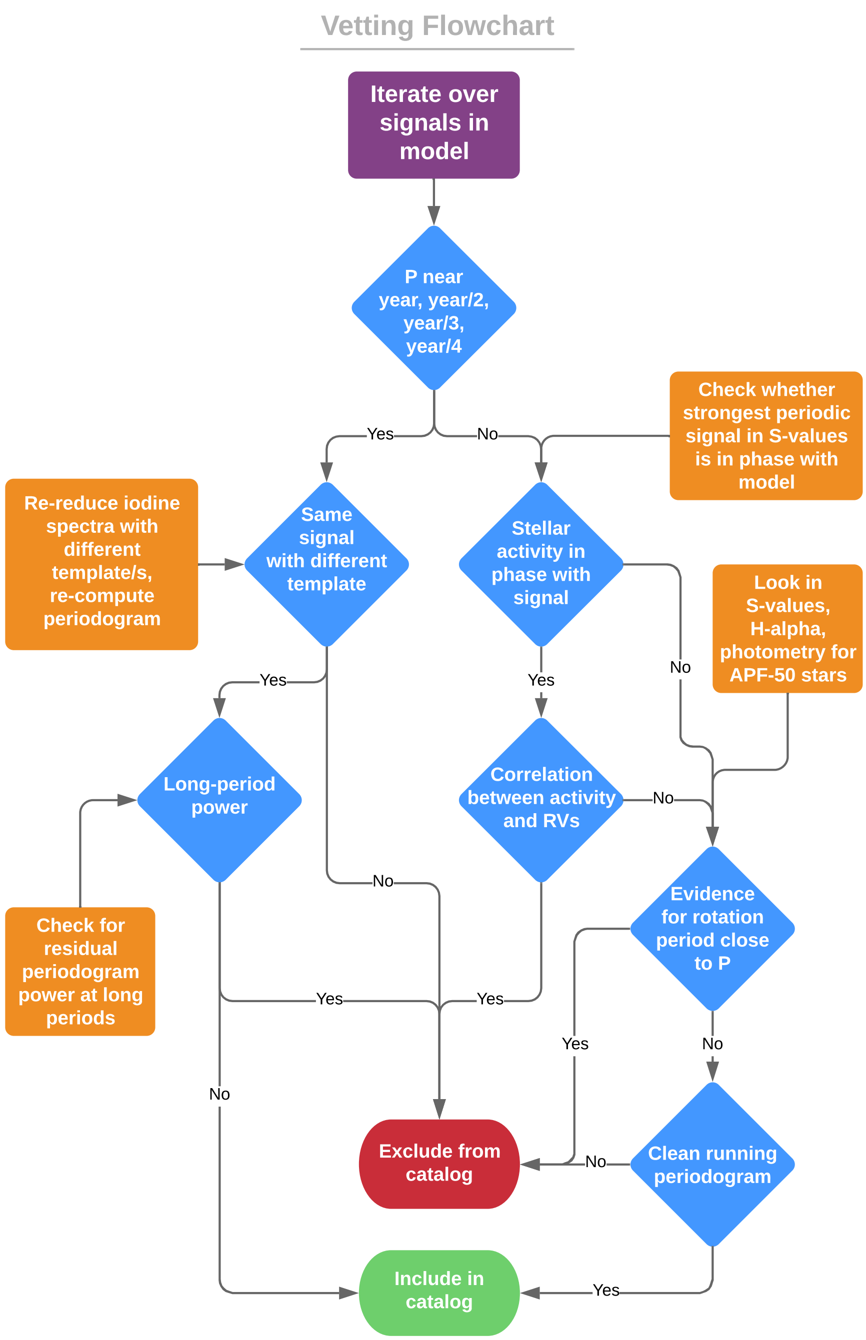

Figure 2.9: Candidate vetting flowchart.

2.6 Planet and Stellar/Substellar Companion Catalog

We present orbital solutions for the known planets, substellar companions, and stellar binaries that `RVSearch` has recovered in the California Legacy Survey. As mentioned in Section 5.1, where appropriate, we modeled long-period companions with linear or parabolic trends. We included in the supplemental information portions of the tables associated with each class of object: one for planets, one for stellar and substellar companions that are best modeled by Keplerian orbits, and one for stars with linear or parabolic RV trends. We also present 14 newly confirmed or significantly revised exoplanets and substellar companions. We list them and their orbital parameters in Table 1, and include individual notes on each system in Appendix A. Figure 2.10 shows all recovered planets in our survey, and distinguishes between known planets and new discoveries.

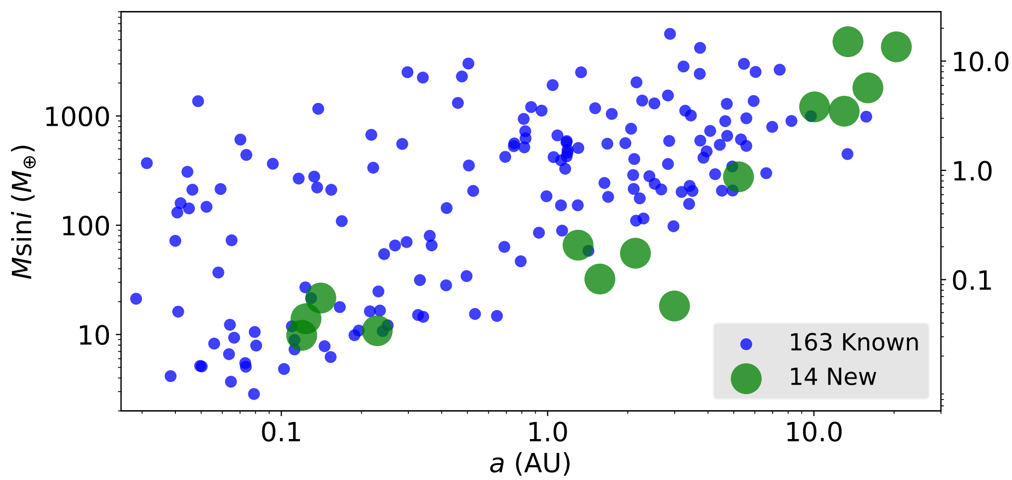

Figure 2.10: Scatterplot of best-fit $M \sin i$ and semi-major axis values for planets in the CLS catalog. Blue dots represent known planets, while green circles represent newly discovered planets and planets with significantly revised orbits.

Table 2.1: Discovered or Revised Planets

Name	$M \sin i$ [M_J] ($M_\oplus$)	a [AU]	e
HD 107148 c	$0.0626^{+0.0097}_{-0.0098}$ ($19.9^{+3.1}_{-3.1}$)	$0.1406^{+0.0018}_{-0.0018}$	$0.34^{+0.13}_{-0.16}$
HD 136925 b	$0.84^{+0.078}_{-0.074}$ (267^{+25}_{-24})	$5.13^{+0.12}_{-0.11}$	$0.103^{+0.094}_{-0.070}$
HD 141004 b	$0.0428^{+0.0047}_{-0.0045}$ ($13.6^{+1.5}_{-1.4}$)	$0.1238^{+0.002}_{-0.002}$	$0.16^{+0.11}_{-0.10}$
HD 145675 c	$5.8^{+1.4}_{-1.0}$	$16.4^{+9.3}_{-4.3}$	$0.45^{+0.17}_{-0.15}$
HD 156668 b	$0.0991^{+0.0079}_{-0.0077}$ ($31.5^{+2.5}_{-2.5}$)	$1.57^{+0.017}_{-0.017}$	$0.089^{+0.084}_{-0.061}$
HD 164922 e	$0.0331^{+0.0031}_{-0.0031}$ ($10.52^{+0.99}_{-0.97}$)	$0.2292^{+0.0026}_{-0.0027}$	$0.086^{+0.083}_{-0.060}$

Table 2.1: Discovered or Revised Planets (Continued)

Name	$M \sin i$ $[M_J]$ $(M_\oplus)$	a [AU]	e
HD 168009 b	$0.03^{+0.0038}_{-0.0037}$ $(9.5^{+1.2}_{-1.2})$	$0.1192^{+0.0017}_{-0.0018}$	$0.121^{+0.110}_{-0.082}$
HD 213472 b	$3.48^{+1.10}_{-0.59}$	$13.0^{+5.7}_{-2.6}$	$0.53^{+0.120}_{-0.085}$
HD 24040 c	$0.201^{+0.027}_{-0.027}$ $(63.9^{+8.6}_{-8.6})$	$1.3^{+0.021}_{-0.021}$	$0.11^{+0.120}_{-0.079}$
HD 26161 b	$13.5^{+8.5}_{-3.7}$	$20.4^{+7.9}_{-4.9}$	$0.82^{+0.061}_{-0.050}$
HD 3765 b	$0.173^{+0.014}_{-0.013}$ $(54.8^{+4.3}_{-4.2})$	$2.108^{+0.032}_{-0.033}$	$0.298^{+0.078}_{-0.071}$
HD 66428 c	27^{+22}_{-17}	$23.0^{+19.0}_{-7.6}$	$0.32^{+0.23}_{-0.16}$
HD 68988 c	$15.0^{+2.8}_{-1.5}$	$13.2^{+5.3}_{-2.0}$	$0.45^{+0.130}_{-0.081}$
HD 95735 c	$0.0568^{+0.0091}_{-0.0083}$ $(18.0^{+2.9}_{-2.6})$	$3.1^{+0.13}_{-0.11}$	$0.14^{+0.160}_{-0.095}$

2.7 Discussion

Through the use of high-cadence APF observations and long-baseline HIRES observations, we have expanded the population of known exoplanets along the current mass and semi-major axis boundary of detectability, as seen in Figure 2.10. We recovered 43 planets with $M \sin i < 30\ M_\oplus$, including four new discoveries within 1 AU. In a future chapter in the California Legacy Survey series, we will leverage the decades-long-baseline datasets in which these planets were discovered, in order to constrain the probability that a host of a small planet also hosts an outer companion, as explored in Bryan, Knutson, Lee, et al., 2019 and Zhu and Wu, 2018. We will also directly place a lower limit on the conditional occurrence of inner small planets given the presence of an outer gas giant.

In addition to expanding the population of small planets with measured $M \sin i$, we discovered or revised the orbits of ten planets with orbital separations greater than 1 AU, six of them beyond 4 AU. We represent the model posteriors for the coldest of these planets in Figure 2.11, and show a gallery of some of their orbits in Figure 2.12. These discoveries include two new detections with incomplete orbits, HD 213472 b and HD 26161 b. Details are provided in Appendices A.3 and A.14. Using HIRES to extend the observational baseline of our survey by another decade will tighten our $M \sin i$ and orbital parameter constraints for these planets, and may reveal more cold companions beyond 10 AU.

In a future chapter in the CLS series, we will use our sample of long-period planets and completeness contours to measure the mass–period planet occurrence distribution out to 10 AU, extending beyond the Keck Planet Search's limit of 5 AU

(Cumming et al., 2008) and the 9 AU limit of Wittenmyer, Wang, et al., 2020. This will provide novel constraints on planet occurrence beyond the water ice line, resolve the discrepancy between the results of Fernandes et al., 2019 and those of Wittenmyer, Wang, et al., 2020, and provide new insight into planet formation across protoplanetary disks.

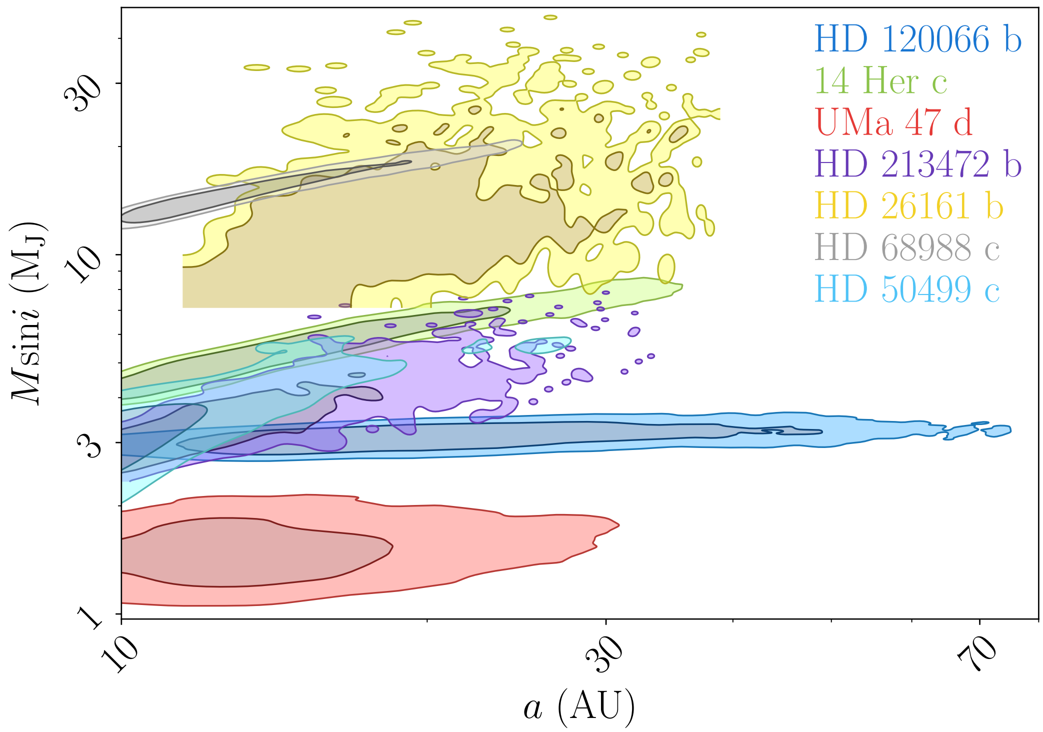

Figure 2.11: Contours (1- and 2-σ) of $M \sin i$ and semi-major axes for planets in the CLS sample whose semi-major axis posteriors extend beyond 10 AU. Contours for HD 26161 b have hard cutoffs due to sparsity below 7 M_J and 12 AU; these limits come from the data's baseline and RV increase to date.

Figure 5.1 is a visualization of the eccentricities of all planets in the California Legacy Survey. In future work, we will quantify the eccentricity distribution of gas giants in our sample and its dependence on planet mass and multiplicity, as well as the eccentricity distributions of brown dwarfs and other substellar companions, in order to clarify possible formation pathways. We will extend the wide-orbit population comparisons of Bowler, Blunt, et al., 2020 to our sample of planets and brown dwarfs within 20 AU of their hosts. We will also explore the eccentricity distribution of gas giants beyond 7 AU. As Figures 2.12 and 5.1 show, all planets recovered beyond 7 AU are eccentric with significance $e > 2\sigma_e$. This may be a selection effect, as the median baseline of observations in our sample is 21 yr, which

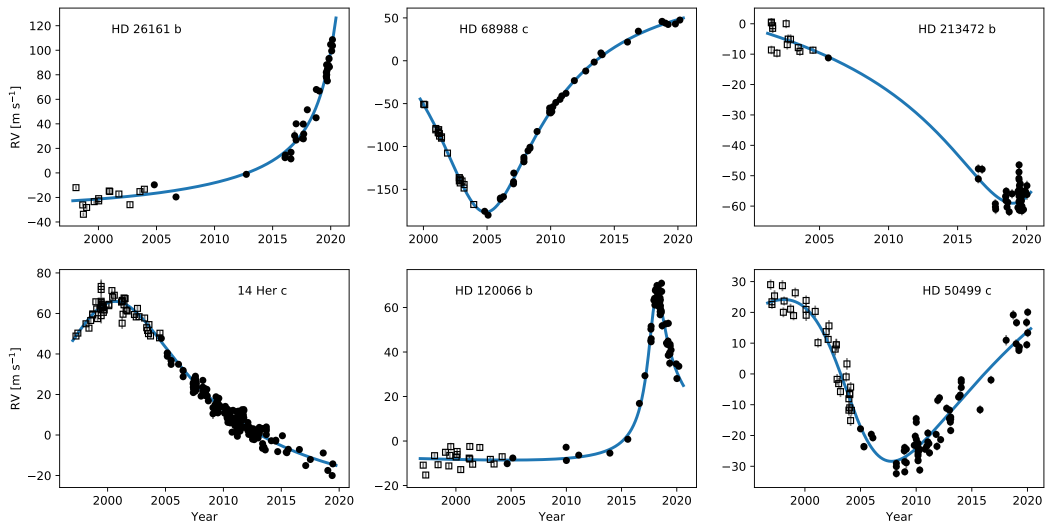

Figure 2.12: Orbit gallery for six of the coldest companions in our survey. We plot RV data and Keplerian model versus year, and subtract off the model signatures of inner companions and stellar activity. We did not include UMa 47 d, seen in Figure 2.11, in this plot, because its detection relied on early Lick-Hamilton RVs, and we wanted to showcase HIRES RV measurements from the past twenty-four years.

corresponds to a semi-major axis of 7.6 AU for a planet orbiting a solar-mass star. It is possible that planets with orbital periods beyond our observational baselines are more easily detectable if they are eccentric. We can use injection-recovery tests to determine whether there is a detection bias toward eccentric planets beyond observational baselines. If this phenomenon is not a selection effect, it might imply that most giant planets beyond 7 AU have undergone a scattering event or otherwise been excited to high eccentricity. Taken together, these studies will leverage this decades-long observational undertaking to provide new insights into planet formation and evolution.

All code, plots, tables, and data used in this chapter are available at `github.com/leerosenthalj/CLSI`. Data and tables, including the full stellar catalog with { M, R, T_{eff}, log$g$, [Fe/H] }, as well as APT photometry, are also available in the associated .tar.gz file available through ApJ. `RVSearch` is available at `github.com/California-Planet-Search/rvsearch`. This research makes use of GNU Parallel (Tange, 2011). We made use of the following publicly available Python modules: `pandas` (McKinney, 2010), `numpy/scipy` (Walt et al., 2011), `emcee` (Foreman-Mackey, Hogg, Lang, et al., 2013b), `Specmatch` (Petigura, 2015; Yee et al., 2017), `Isoclassify` (Huber, 2017), `TheJoker` (Price-Whelan et al., 2017),

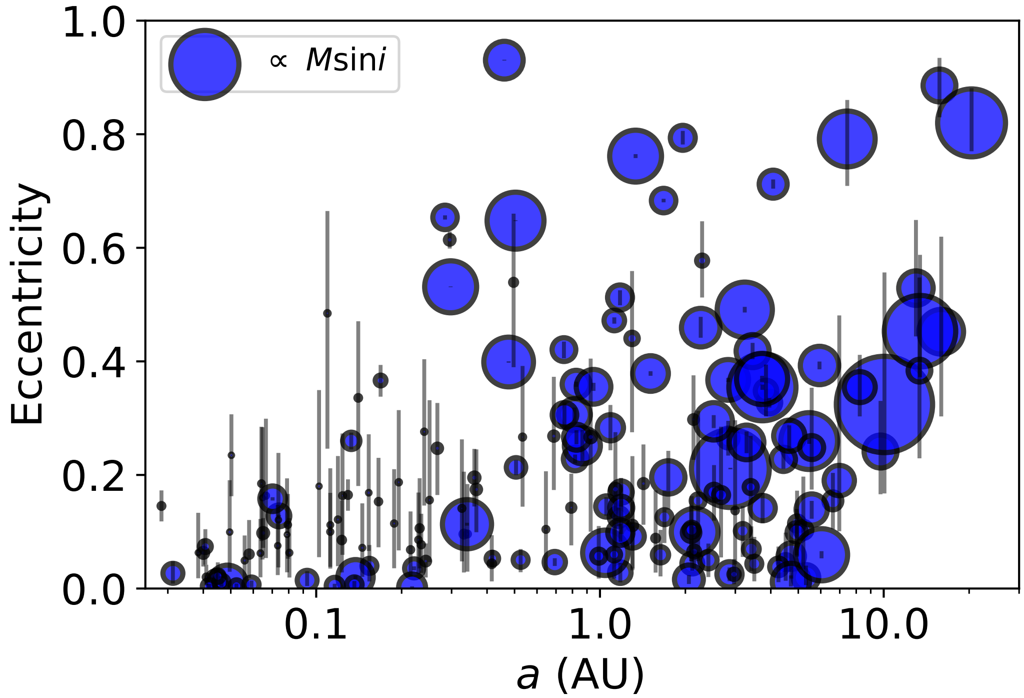

Figure 2.13: $M \sin i$, a, and eccentricity of the CLS sample. Eccentricity is plotted in medians and 68% confidence intervals, while scatter size is proportional to $M \sin i$ posterior mode.

`RadVel` (Fulton, Petigura, Blunt, et al., 2018), `RVSearch` (this work).

2.8 Supplemental Information

As supplemental information, we have included individual notes on each planet discovery reported in this chapter; complete tables of recovered planets, Keplerian-resolved stellar binaries, and substellar companions in the California Legacy Survey; signals that `RVSearch` recovered and we determined to be false-positives; linear and parabolic RV trends; and the stellar sample and RV dataset. Associated tables are available in their entirety at `github.com/leerosenthalj/CLSI`.

Individual Discoveries and Revised Orbits

HD 3765

HD 3765 is a K2 dwarf at a distance of 17.9 pc (Gaia Collaboration, Brown, Vallenari, Prusti, de Bruijne, Babusiaux, et al., 2018). Figure 2.14 shows the `RVSearch` results for this star. We recovered a signal with a period of 3.36 yr. Table 1 reports all planet parameters. There is significant periodicity in the S-value time series, but concentrated around a period of 12 yr. Furthermore, we find no correlation between the RVs and S-values. Thus, we label this signal as a confirmed planet, with $M \sin i = 0.173 \pm 0.014\ M_{\mathrm{J}}$ and $a = 2.108 \pm 0.033$ AU. The magnetic activity cycle is too weak for `RVSearch` to recover, but is evident in the best-fit RV residuals. We used RadVel to model this activity cycle with a squared-exponential Gaussian process, and report MCMC-generated posteriors for both orbital and Gaussian process parameters in Figure 2.16 and Figure 2.17.

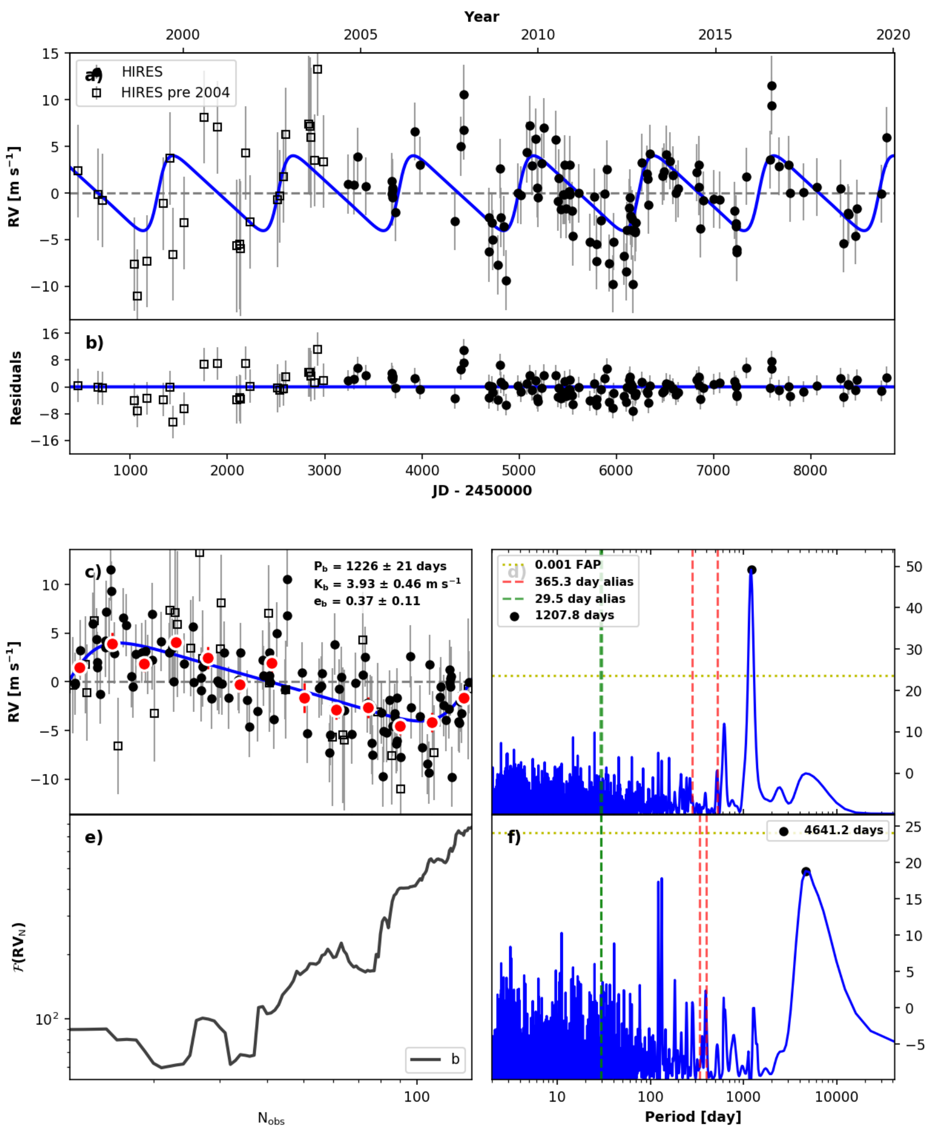

Figure 2.14: `RVSearch` summary plot for HD 3765. See Figure 2.6 for plot description.

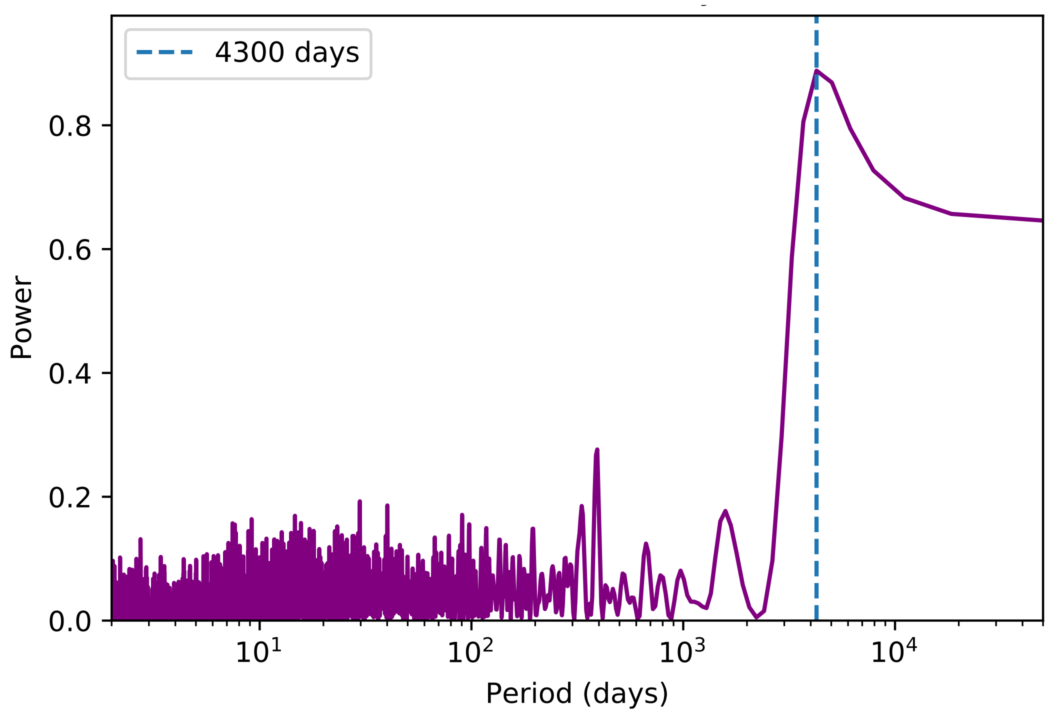

Figure 2.15: Lomb-Scargle periodogram of HIRES S-values for HD 3765. Significant power at and beyond 4,300 days.

HD 24040

HD 24040 is a G1 dwarf at a distance of 46.7 pc. Figure 2.18 shows the `RVSearch` results for this star. It hosts a known gas giant (Wright, Marcy, Fischer, et al., 2007; Feng et al., 2015) with a semi-major axis that we measured as $a = 4.72 \pm 0.18$ AU, an orbital period of 9.53 ± 10^{-4} years, and a minimum mass $M \sin i = 4.09 \pm 0.22\ M_J$. We have extended the observational baseline of our HIRES measurements to 21.7 years, constrained the long-term trend and curvature of the RVs, and discovered a new exoplanet, a sub-Saturn ($M \sin i = 0.201 \pm 0.027\ M_J$) on a 1.4 yr orbit ($a = 1.30 \pm 0.021$ AU) that is consistent with circular. The S-values are uncorrelated with the the RVs of both planet signals, after removing the long-term trend. Table 1 reports all planet parameters.

In addition to the newly detected sub-Saturn, we further constrained the known linear trend in the RVs, and found evidence for a curvature term as well. `RVSearch` detected a curvature term with model preference ΔBIC> 10 over a purely linear trend. We measured the linear trend to be 0.00581 ± 0.00044 m s^{-1} yr^{-1}, and the curvature to be $-6.6 \times 10^{-7} \pm 1.2 \times 10^{-7}$ m s^{-1} yr^{-1}, a 5.5-σ detection. The trend and

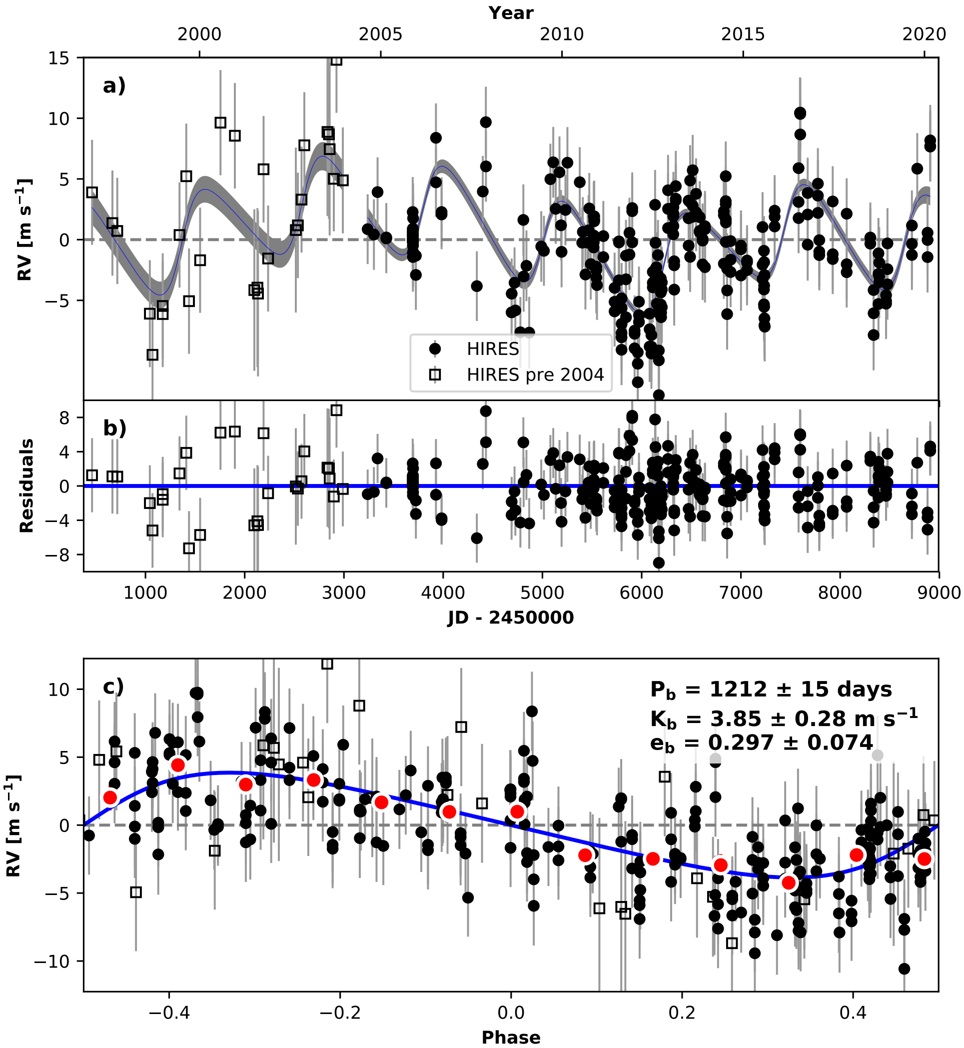

Figure 2.16: `RadVel` model orbital plot for HD 3765, including a Gaussian process with a squared-exponential kernel. The grey shaded curve represents the 68% interval for the Gaussian process RV signature.

curvature parameters are slightly correlated in the posterior, but neither is correlated with any of the Keplerian orbital parameters in the model. Therefore, we kept the curvature term that `RVSearch` selected in our model. This long-term trend is low-amplitude enough that it may be caused by another planet in the system, orbiting beyond 30 AU. Gaia astrometry or another two decades of RVs may provide further constraints on this object.

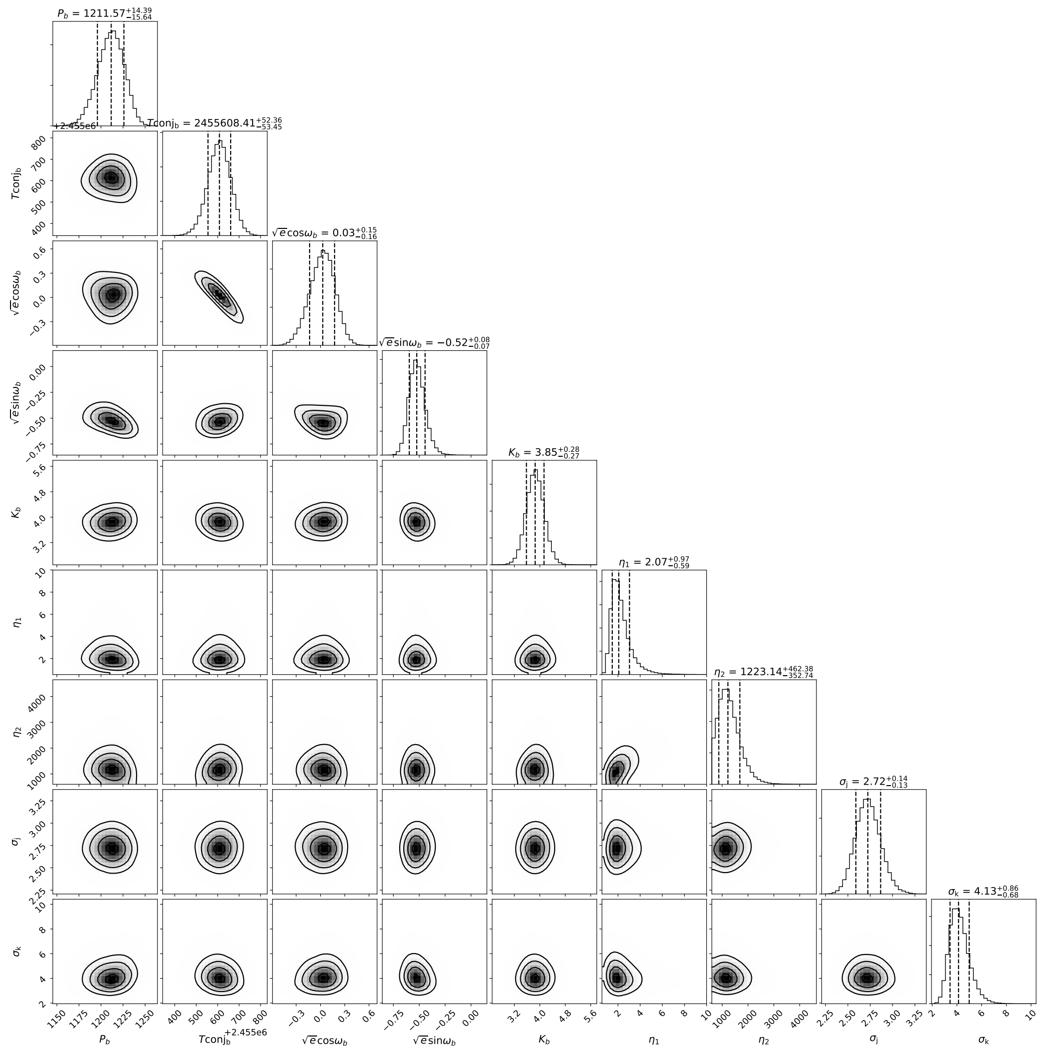

Figure 2.17: Orbital and Gaussian process parameter posteriors for HD 3765. η_1 is the GP amplitude, while η_2 is the GP exponential decay timescale.

HD 26161

HD 26161 is a G0 dwarf located at a distance of 50.0 pc. Figure 2.20 shows the `RVSearch` results for this star. Our RVs are consistent with a long-period, eccentric companion, and `RVSearch` detected this long-period signal. Due to the sparseness of the data and the fractional orbital coverage, traditional MCMC methods fail to return a well-sampled model posterior. Since the data underconstrains our model, we used `TheJoker` to sample the posterior, which is consistent with an extremely long-period gas giant with minimum mass $M \sin i = 13.5^{+8.5}_{-3.7}$ M_J, semi-major axis $a = 20.4^{+7.9}_{-4.9}$ AU, and eccentricity $e = 0.82^{+0.06}_{-0.05}$. Table 1 reports current estimates

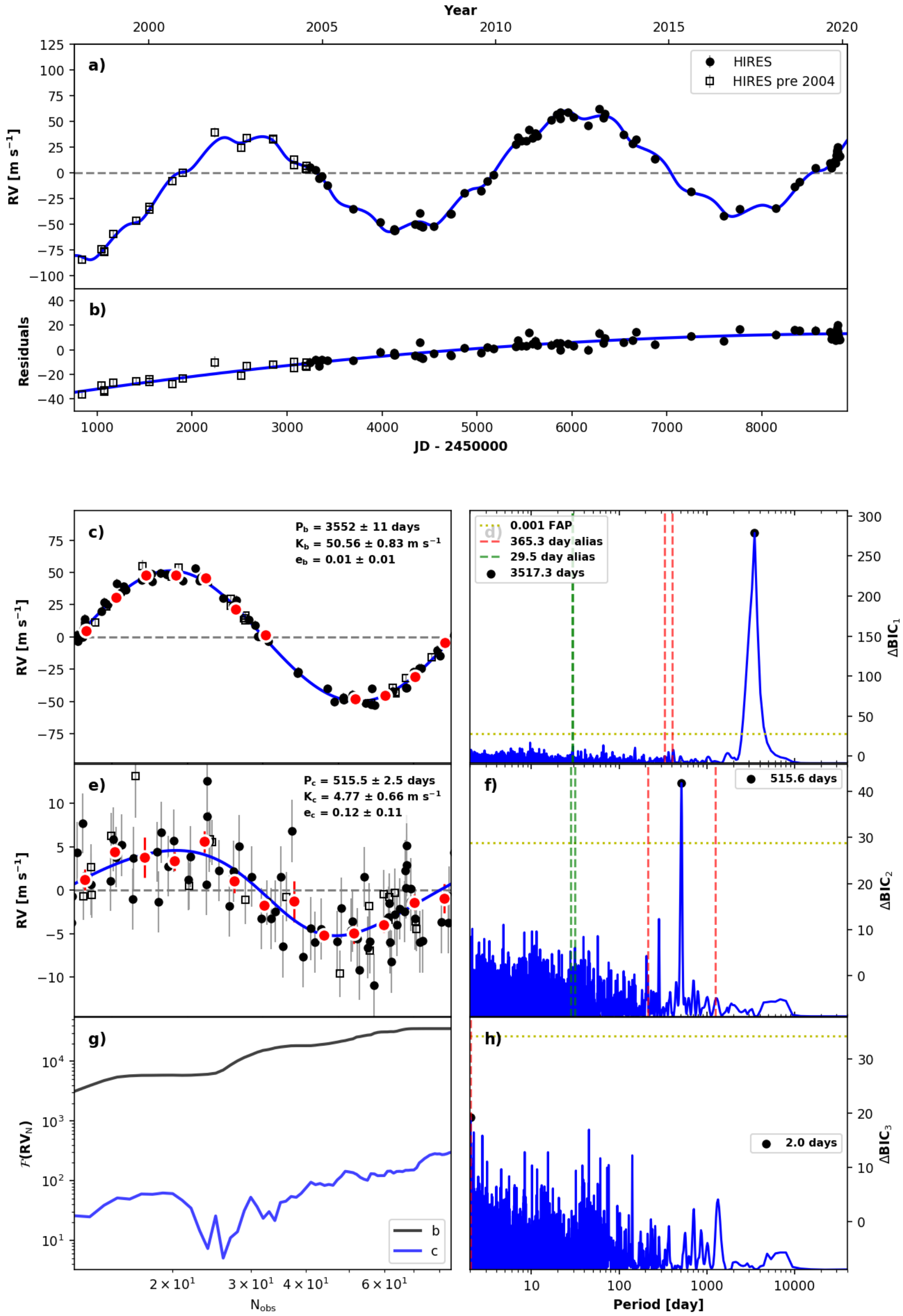

Figure 2.18: `RVSearch` summary plot for HD 24040. See Figure 2.6 for plot description.

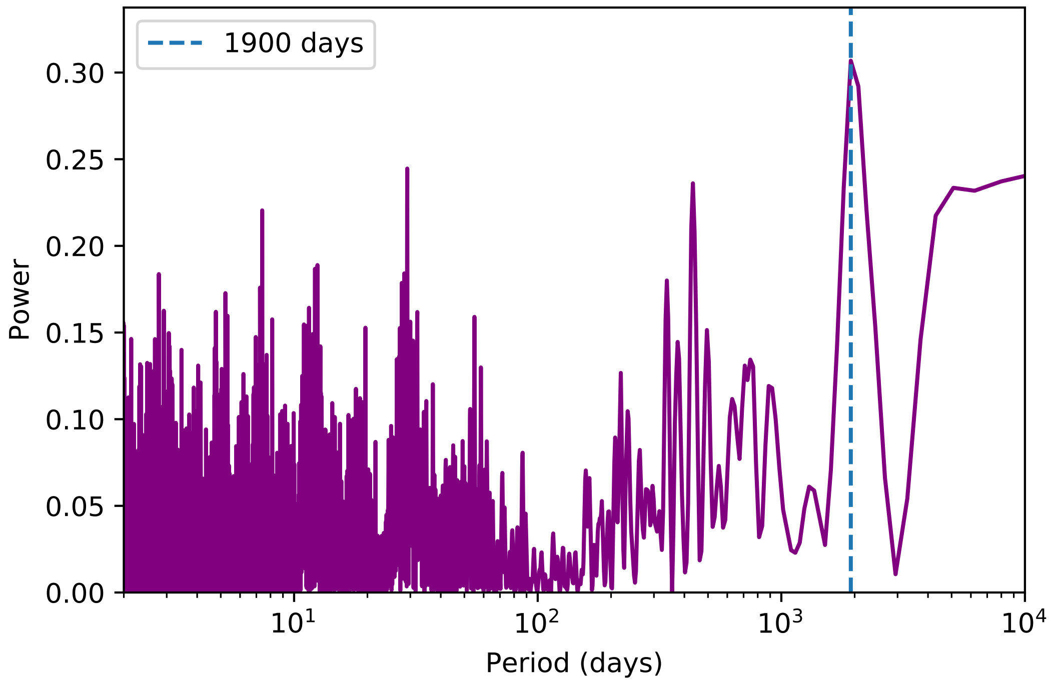

Figure 2.19: Lomb-Scargle periodogram of HIRES S-values for HD 24040. No periods show power that is statistically significant.

of all orbital parameters, and Figure 2.21 shows their posterior distributions. A Keplerian model is significantly preferred over a quadratic trend, with ΔBIC> 15.

The Simbad stellar catalog designates HD 26161 as a stellar multiple. We used Gaia to identify a binary companion with similar parallax and within 60 arcseconds. This companion has an effective temperature identified from Gaia colors of 4,053 K, and a projected separation of 562 AU. A stellar companion that is currently separated from its primary by more than 560 AU could not cause a change in radial velocity of 100 m/s over 4 years. This curve is far more likely caused by an inner planetary or substellar companion approaching periastron.

Figure 2.22 shows a sample of possible orbits for HD 26161 b, drawn from our rejection sampling posteriors and projected over the next decade. We will continue to monitor HD 26161 with HIRES at moderate cadence, and have begun observing this star with APF. As we gather more data during the approach to periastron, we can tighten our constraints on the minimum mass, eccentricity, and orbital separation of HD 26161 b.

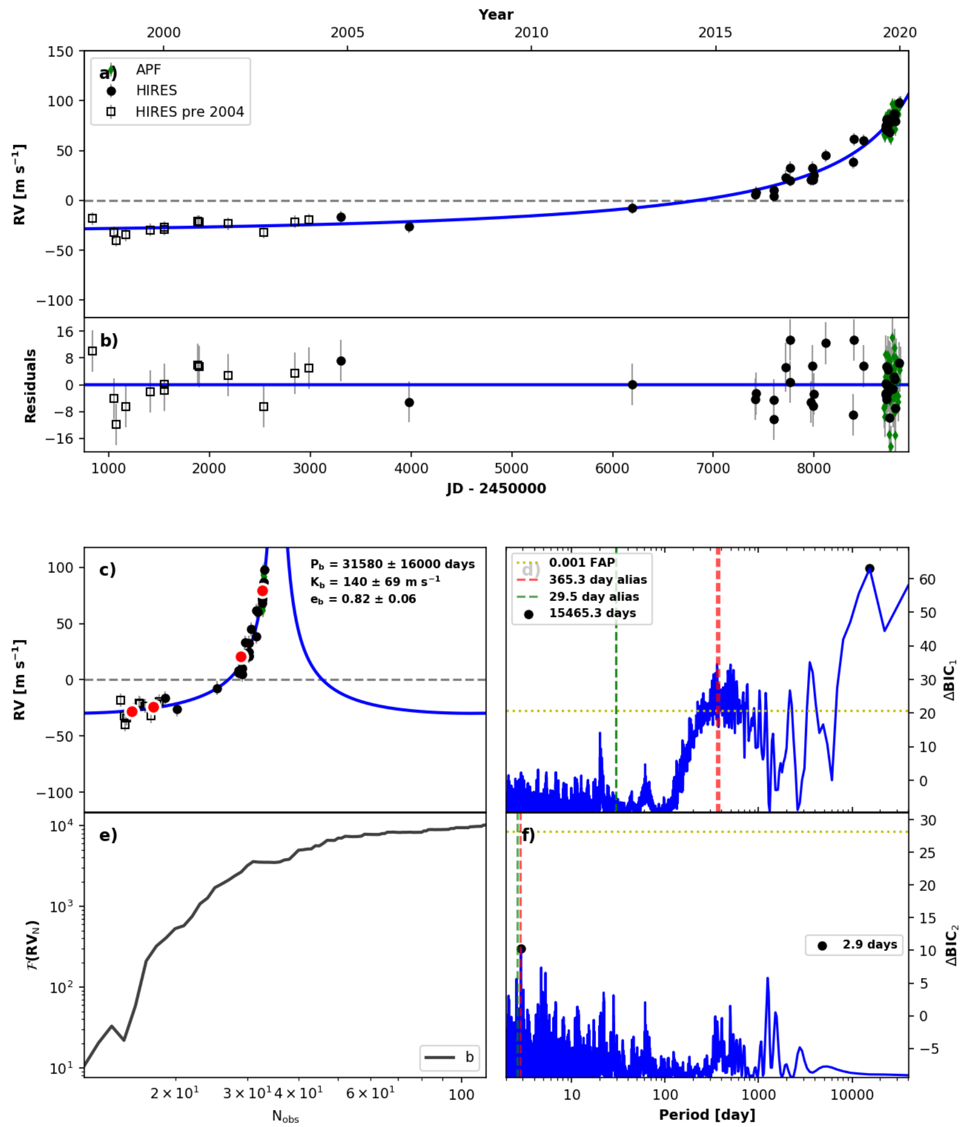

Figure 2.20: `RVSearch` summary plot for HD 26161. See Figure 2.6 for plot description.

HD 66428

HD 66428 is a G8 dwarf found at a distance of 53.4 pc. Figure 2.23 shows the `RVSearch` results for this star. This system has one well-constrained cold Jupiter (Butler, Wright, et al., 2006) and an outer companion candidate first characterized in Bryan, Knutson, Howard, et al., 2016 as a linear trend. With four more years of HIRES data, we now see curvature in the RVs and a clear detection in `RVSearch`, and can place constraints on this outer candidate's orbit with a Keplerian model. The

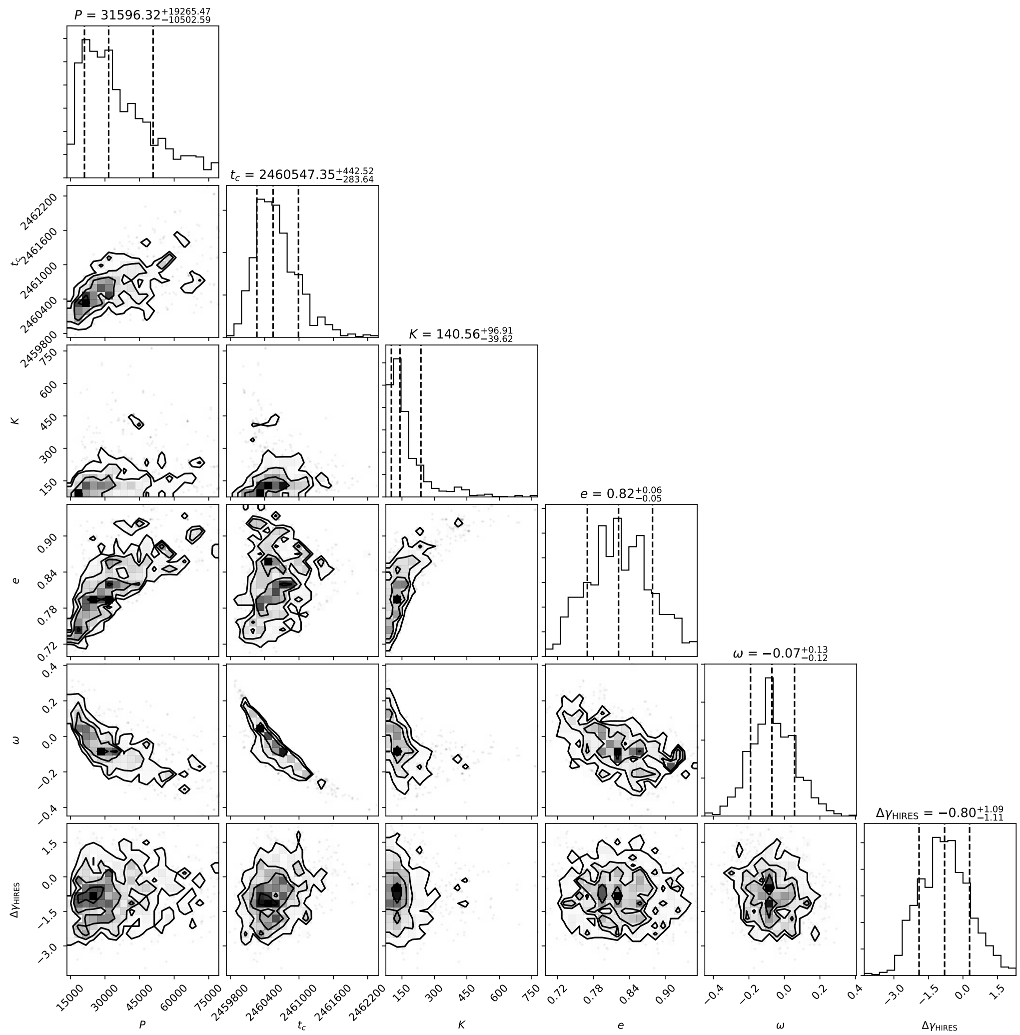

Figure 2.21: Rejection sampling posterior for HD 26161.

Keplerian orbit for the outer candidate is preferred to a parabolic trend with ΔBIC> 30. A maximum likelihood fit gives an orbital period of $P = 36.4$ yr. However, since we have only observed a partially resolved orbit so far, the orbit posterior in period-space is wide and asymmetric. MCMC sampling produces $P = 88^{+153}_{-49}$ yr. Table 1 reports current estimates of all orbital parameters.

The model parameters are $M \sin i = 27^{+22}_{-17}\ M_{\mathrm{J}}$, $a = 23.0^{+19.0}_{-7.6}$ AU, and $e = 0.31^{+0.13}_{-0.13}$. This orbital companion could be a massive gas giant or a low-mass star, if we only consider constraints from RV modeling. However, Bryan, Knutson, Howard, et al., 2016 used NIRC2 Adaptive-Optics images to place upper bounds on the mass and semi-major axis of an outer companion, at a time when it only presented as a linear

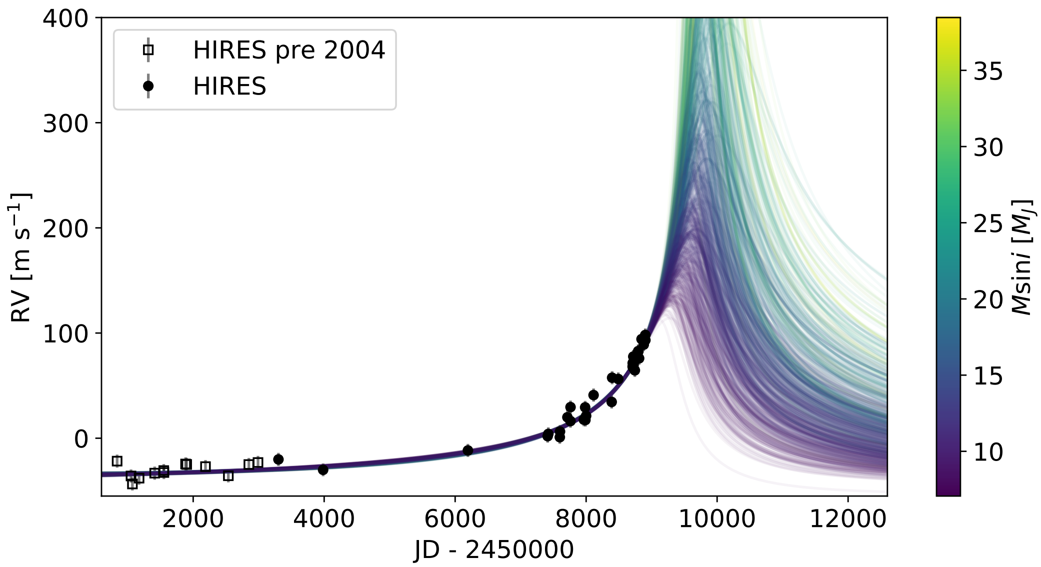

Figure 2.22: Possible orbits for HD 26161 b. RV curves are drawn from the rejection sampling posterior generated with `TheJoker`. The color of each orbit drawn from the posterior scales with $M \sin i$.

trend in HIRES RVs. They found an upper bound of ≈100 M_J on mass, not just $M \sin i$, and an upper bound of ≈150 AU on a. We will continue to monitor this star with HIRES to further constrain the mass and orbit of HD 66428 c.

HD 68988

HD 68988 is a G0 dwarf found at a distance of 61 pc. Figure 2.24 shows the `RVSearch` results for this star. This system has one well-constrained hot Jupiter (Vogt, Butler, Marcy, et al., 2002) and an outer companion candidate that was first characterized in Bryan, Knutson, Howard, et al., 2016 as a partially resolved Keplerian orbit. With four more years of HIRES data, we can place tighter constraints on this outer candidate's orbit. A maximum likelihood fit gives an orbital period of 49.2 yr. However, since we have only observed a partially resolved orbit so far, the orbit posterior is wide and asymmetric in period space. MCMC sampling produces $P = 61^{+28}_{-20}$ yr. The model parameters are $M \sin i = 17.6^{+2.4}_{-2.5}$ M_J, $a = 16.5^{+4.8}_{-3.8}$ AU, and $e = 0.53^{+0.13}_{-0.09}$. Table 1 reports all companion parameters.

`RVSearch` detects a third periodic signal, with P = 1,900 days, that has the same period and phase as the peak period in the S-value time-series. This signal also has a low RV amplitude, 6 m s^{-1}. Therefore, we designated this signal as a false-positive

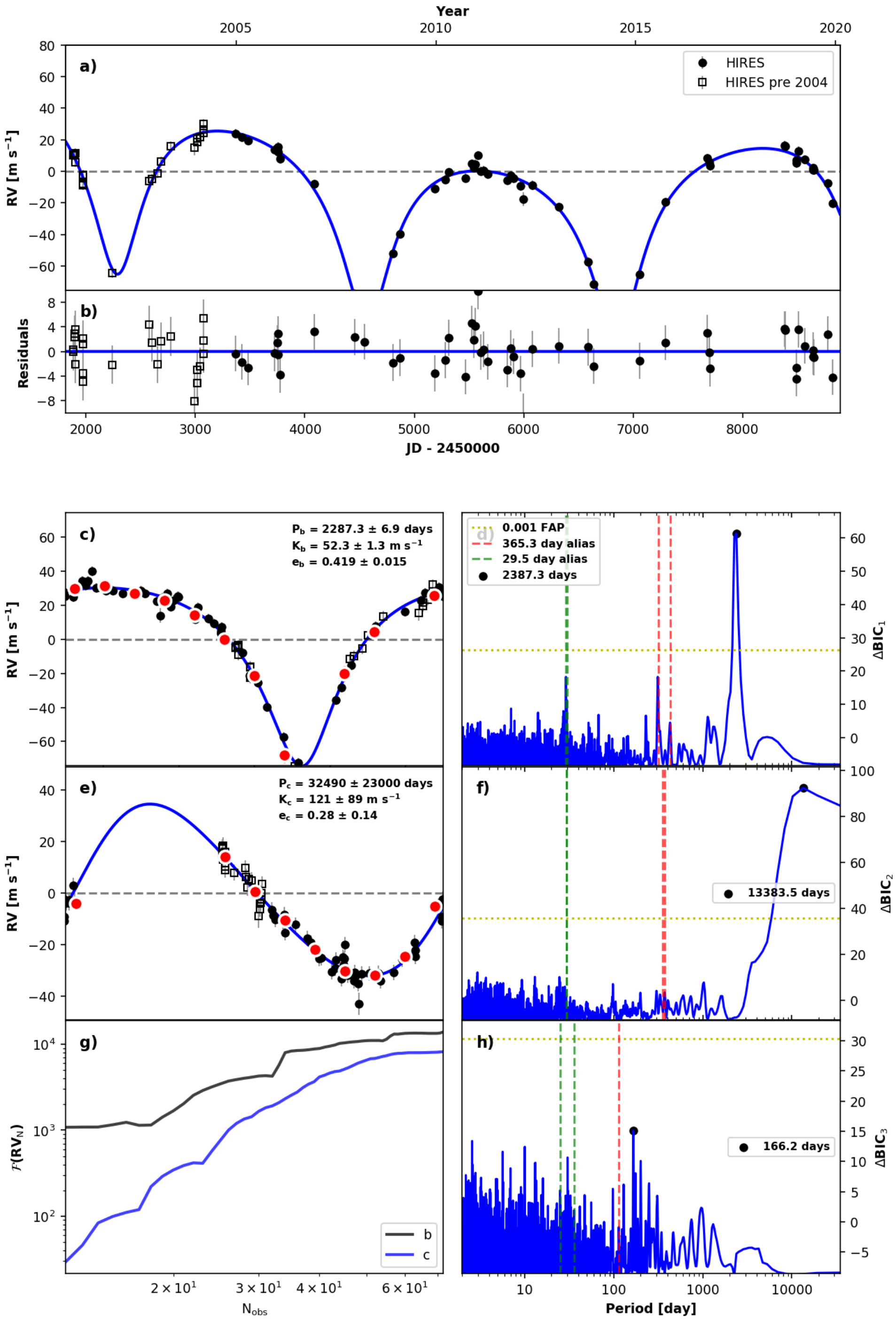

Figure 2.23: RVSearch summary plot for HD 66428. See Figure 2.6 for plot description.

corresponding to stellar activity.

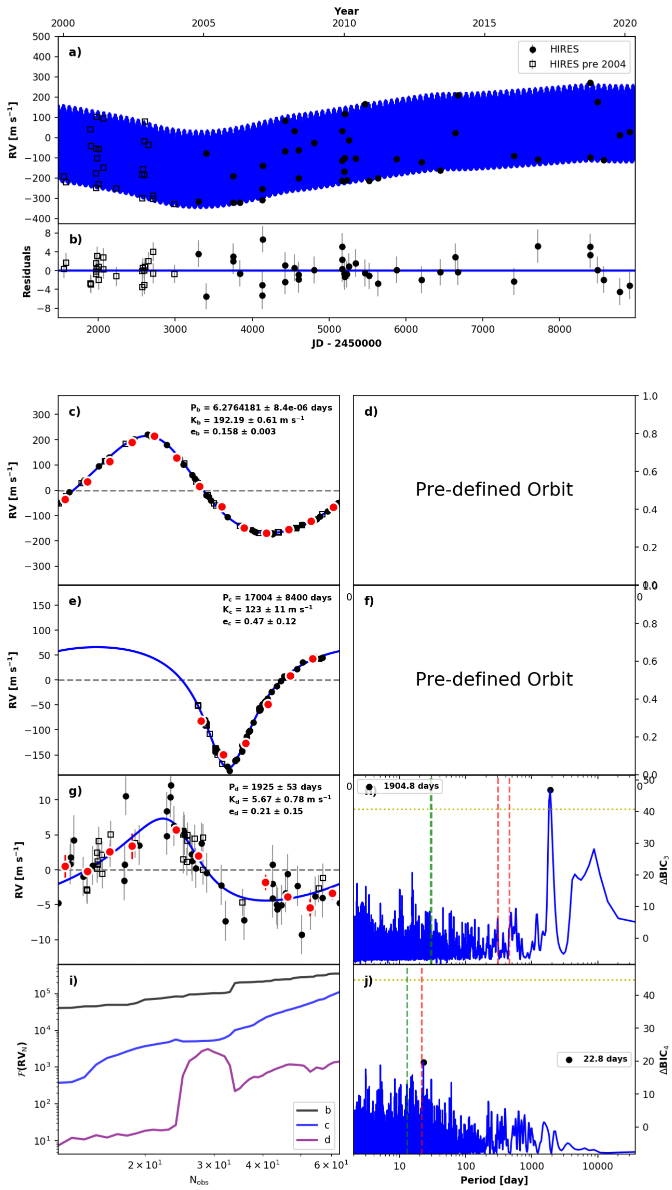

Figure 2.24: RVSearch summary plot for HD 68988. See Figure 2.6 for plot description.

HD 95735

HD 95735 (GJ 411) is an M2 dwarf found at a distance of 2.55 pc. Figure 2.25 shows the RVSearch results for this star. This system has one known short-period super-Earth, with $M \sin i = 3.53\ M_\oplus$ and an orbital period of 12.9 d. Our detection of this planet was driven by high-cadence APF data. This planet was first reported in Díaz et al., 2019, which also noted long-period power in their SOPHIE RV data, but did not have a sufficiently long baseline or the activity metrics necessary to determine the origin of this power. With our HIRES post-upgrade and APF observations, we have an observational baseline of 14 years, allowing us to confirm this long-period signal as a planet with $M \sin i = 24.7 \pm 3.6\ M_{\oplus}$ and an orbital period $P = 8.46$ yr. Table 1 reports all planet parameters. Since GJ 4ll is a cool M dwarf, the Lick-Hamilton and HIRES pre-upgrade data are not reliable, because those detectors are not sufficiently high-resolution to capture a cool M dwarf's dense spectral lines (Fischer, Marcy, et al., 2014).

There is a long-period trend in the HIRES S-value time series, with significant power at and beyond 25 years, but no significant power near the orbital period of the outer candidate. Therefore, we included this candidate in our catalog as a new planet candidate, to be verified and constrained with several more years of HIRES observations.

`RVSearch` also recovered a highly eccentric, 216 day signal, but this signal correlates with APF systematics. Therefore, we labeled it as a false positive. This systematic remained when we applied `RVSearch` only to the HIRES post-upgrade and APF data, and left out the problematic pre-upgrade and Lick data.

HD 107148

HD 107148 is a G5 dwarf at a distance of 49.5 pc. Figure 2.27 shows the `RVSearch` results for this star. Butler, Wright, et al., 2006 reported a planet with a period of 44 days; they reported periodicity at 77 days, but determined that this was an alias of the 44-day signal. The 77-day signal is significantly stronger in our likelihood periodogram, as seen in Figure 2.27, and better fits the data than a 44-day Keplerian by a significant ΔBIC. This constitutes strong evidence that the true period of this planet is 77 days. We report new orbital parameters for this planet in Table 3.

We also recovered a signal with a period of 18.3 days. There is significant periodicity in the S-value time series, a periodogram of which is shown in 2.28, but it concentrated around a period of 6 yr, and there is no significant power near 18.3 days. Furthermore, we find no correlation between the RVs and S-values. Thus, we report this signal as a confirmed planet, with $M \sin i = 19.9 \pm 3.1\ M_{\oplus}$ and $a = 0.1406 \pm 0.0018$ AU.

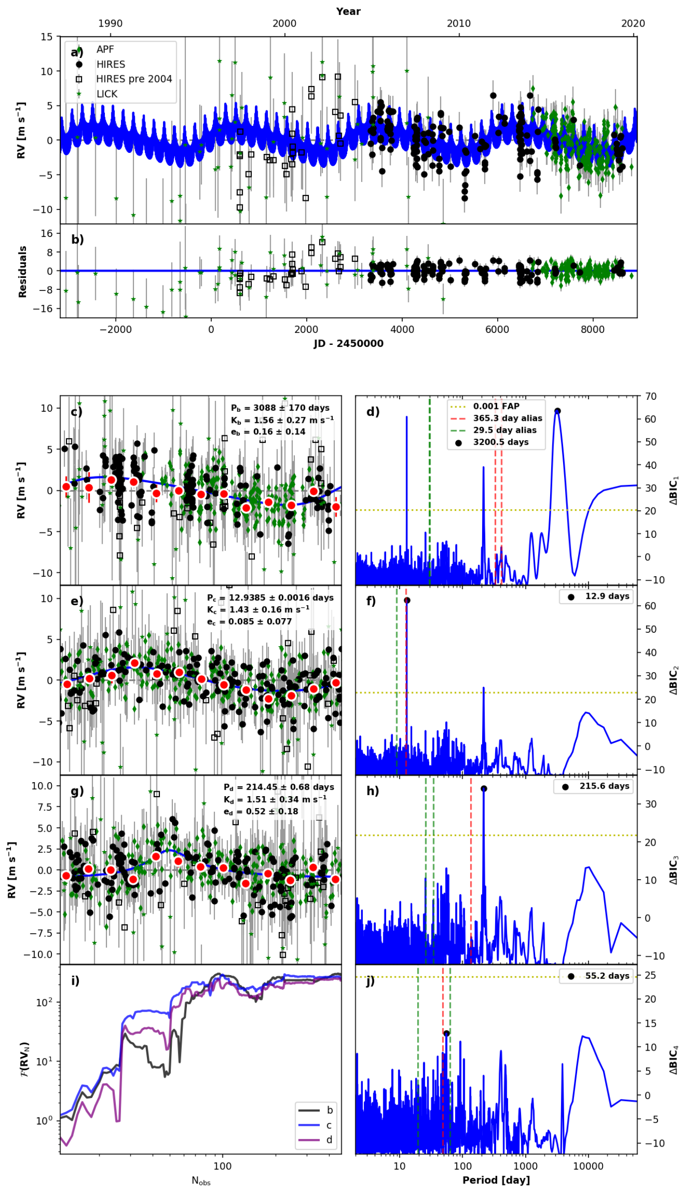

Figure 2.25: `RVSearch` summary plot for HD 95735. See Figure 2.6 for plot description.

HD 136925

HD 136925 is a G0 dwarf, found at a distance of 47.9 pc. `RVSearch` detected two periodic signals in this dataset, as seen in Figure 2.29, at 311 days and 12.4 years. This dataset is currently sparse, with two gaps of several years in the post-upgrade HIRES data, but there is clear long-period variation in the RVs. Keplerian modeling

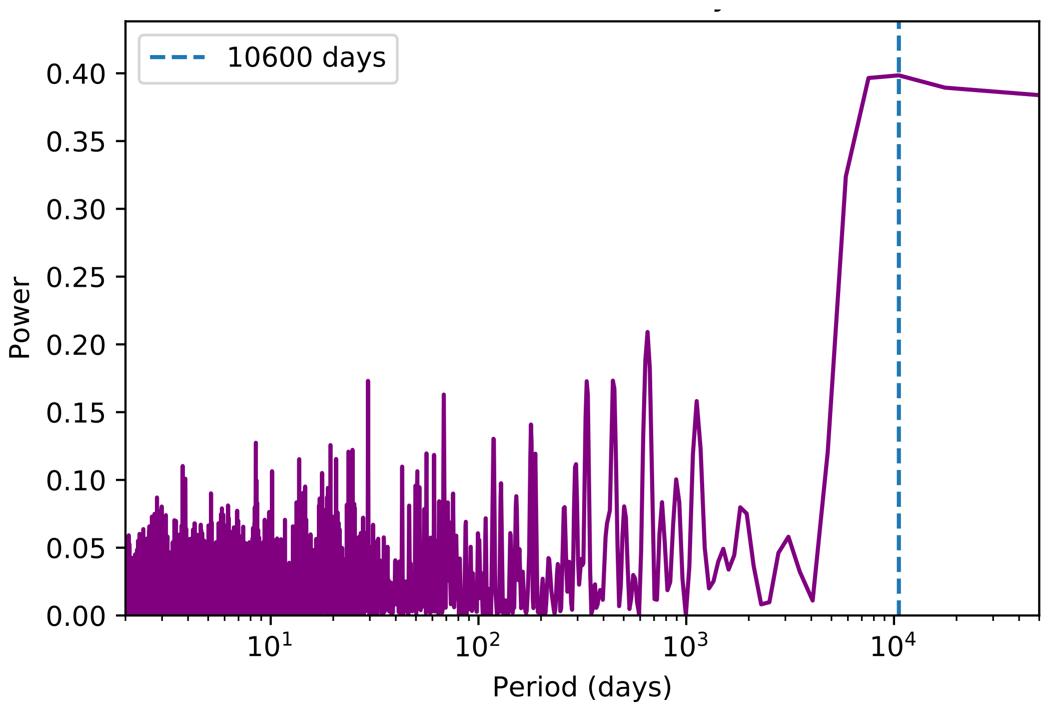

Figure 2.26: Lomb-Scargle periodogram of HIRES S-values for HD 95735. There is evidence for an activity cycle longer than 10,000 days, but no significant power near the period of our 3,000-day planet candidate.

predicts $M \sin i = 0.84\ M_J$ for the giant planet.

The S-value periodogram seen in Figure 2.30 shows no significant power beyond 1,000 days, suggesting that the long-period HD 136925 b is a real planet. There is broad power around 300 days, overlapping with the period of the inner signal. It is unclear whether this periodicity is caused by real stellar variability, or a product of sparse data. Table 1 reports current estimates of all planet parameters. We need more data in order to clarify our model, and determine whether the inner signal is caused by a planet or a product of stellar activity and sparse data. Therefore, we designated HD 136925 b as a planet, and the inner signal as a probable false-positive, to be clarified with continued HIRES observing.

HD 141004

HD 141004 is a G0 dwarf found at a distance of 11.8 pc. Figure 2.31 shows the RVSearch results for this star. `RVSearch` discovered a sub-Neptune at an orbital period of 15.5 days, with $M \sin i = 13.9 \pm 1.5\ M_\oplus$. Table 1 reports current estimates of all planet parameters.

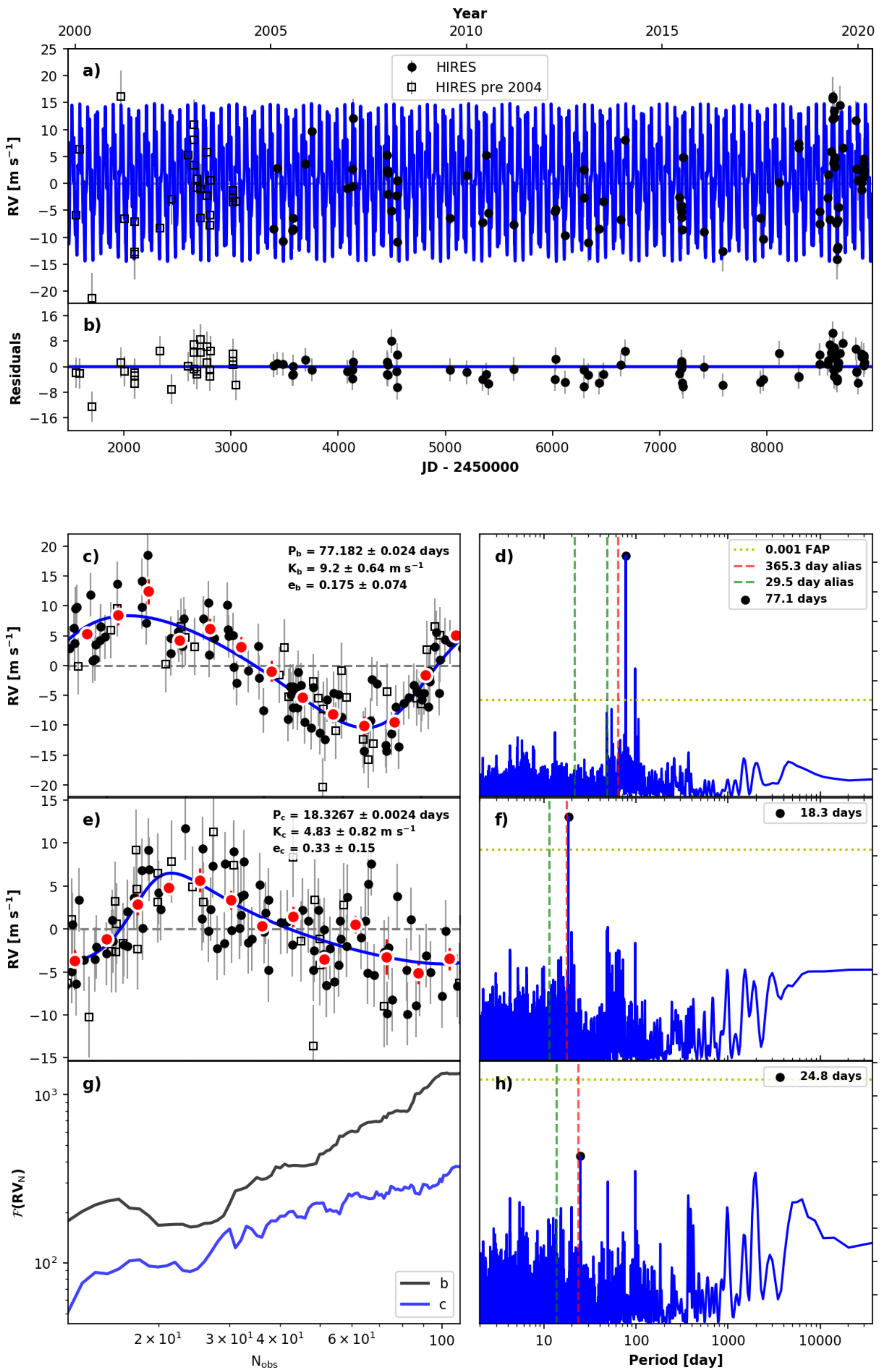

Figure 2.27: RVSearch summary plot for HD 107148. See Figure 2.6 for plot description.

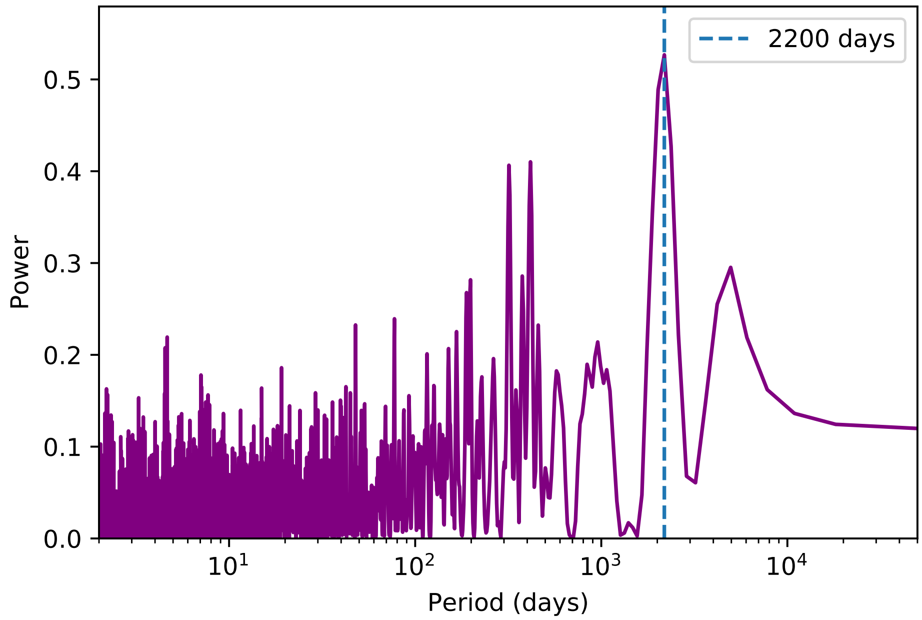

Figure 2.28: Lomb-Scargle periodogram of HIRES S-values for HD 107148. Significant power at and beyond 4,300 days.

HD 145675

HD 145675 (14 Her) is a K0 dwarf found at a distance of 17.9 pc. Figure 2.32 shows the `RVSearch` results for this star. This system has one known cold gas giant, with $M \sin i$= 5.10 M_{J} and an orbital period of 4.84 yr, which was first reported in Butler, Marcy, Vogt, et al., 2003. Wittenmyer, Endl, and Cochran, 2007 conducted further analysis with a longer observational baseline of twelve years, and noted a long-period trend. Wright, Marcy, Fischer, et al., 2007 used additional RV curvature constraints to show that this trend must correspond to a companion with $P > 12$ yr and $M \sin i > 5$ M_{J}. The observational baseline has since increased from twelve years to 22, and regular observations with HIRES and APF allow us to place further constraints on this long-period companion. We find $M \sin i = 5.8^{+1.4}_{-1.0}$ M_{J}, $P = 68^{+64}_{-25}$ yr, semi-major axis $a = 16.4^{+9.3}_{-4.3}$ AU, and eccentricity $e = 0.45^{+0.17}_{-0.15}$. Table 1 reports all planet parameters.

There is strong periodicity in the HIRES S-value time series, peaking around 10 yr, but no significant power near the supposed orbital period of the long-period candidate. These S-values strongly correlate with a third Keplerian signal picked up by our search, also with a period of 10 yr, as seen in the Figure 2.34, therefore

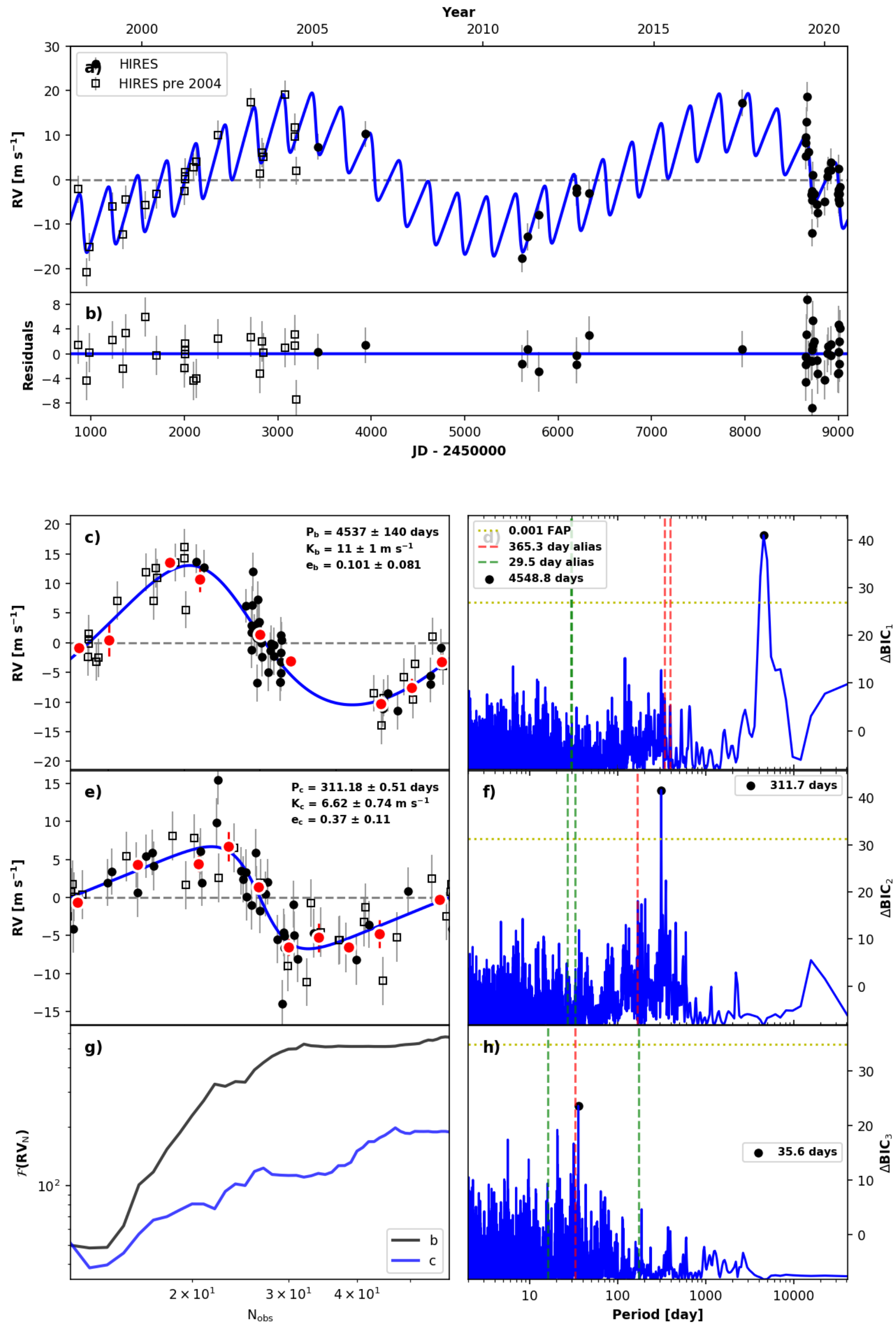

Figure 2.29: `RVSearch` summary plot for HD 136925. See Figure 2.6 for plot description.

we designate this signal as stellar activity.

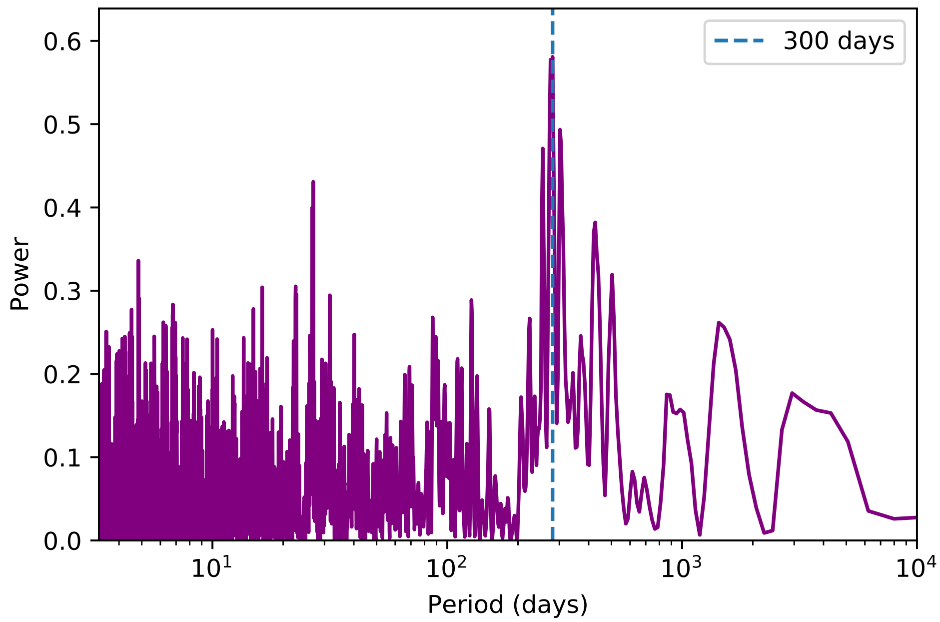

Figure 2.30: Periodogram of HIRES S-values for HD 136925. Significant periodicity around 300 days, near the period of the inner signal.

There is a potential complication owed to a stellar binary candidate. Roberts et al., 2011 conducted a direct-imaging survey of known exoplanet hosts and reported a candidate stellar companion to 14 Her, with a differential magnitude of 10.9 ± 1.0, an angular separation of 4.3", and a minimum orbital separation of 78 AU. This is a single-epoch detection, and therefore could be only a visual binary. Additionally, Rodigas et al., 2011 conducted a deep direct imaging study of 14 Her, to constrain the mass and orbital parameters of 14 Her c, which, at the time, presented only as a parabolic trend in RV data. They used the Clio-2 photometer on the MMT, which has a 9" x 30" field of view; the authors only looked at imaging data within 2", to filter out background stars. Although this deep imaging study did not mention any stellar companion, the candidate reported by Roberts et al., 2011 falls outside of their considered imaging data, which corresponds to a minimum separation of 112.8 AU. Wittrock et al., 2017 also found a null binary detection, using the Differential Speckle Survey Instrument (DSSI) at the Gemini-North Observatory. A 6 Jupiter mass object would not have been detected by the above surveys, as they were designed only to rule out stellar companions, and therefore used shorter imaging exposures that would miss planetary-mass companions.

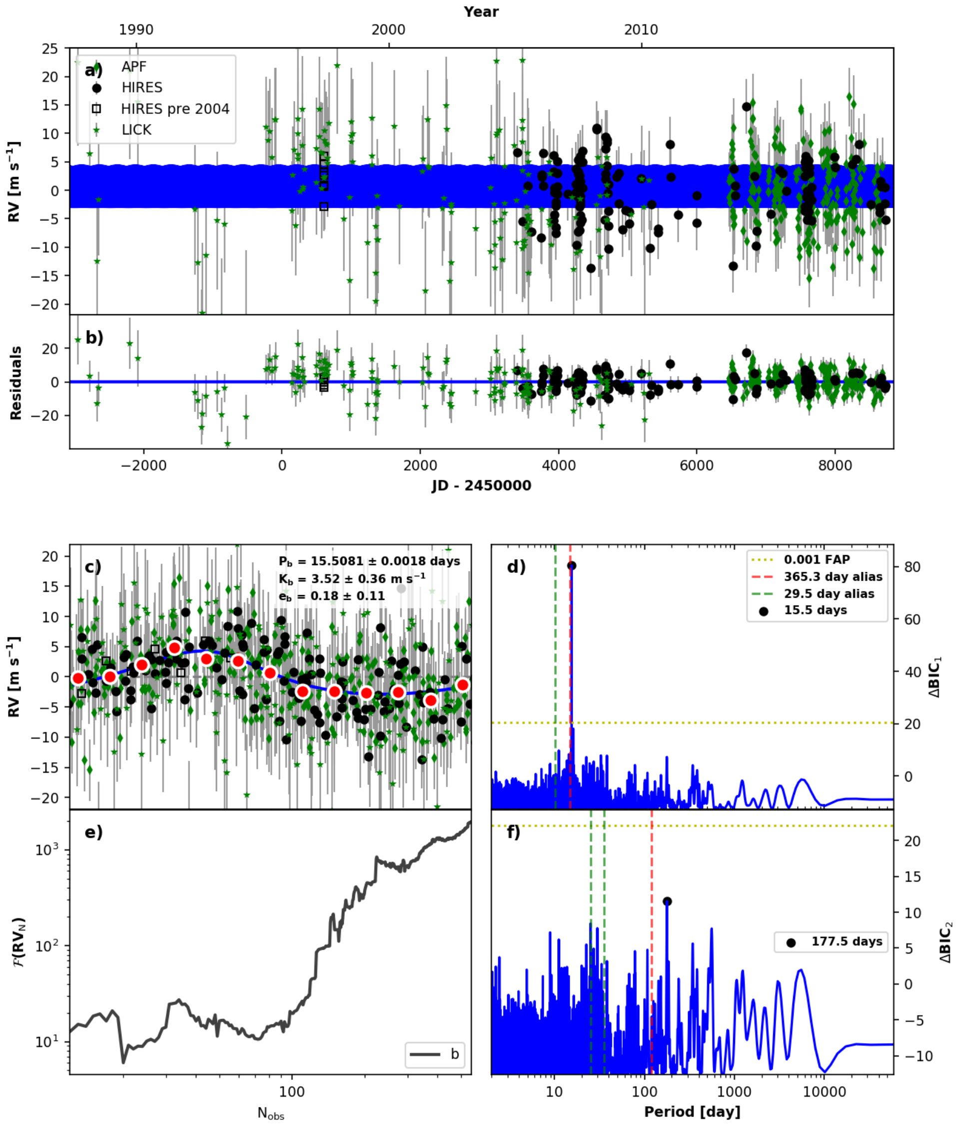

Figure 2.31: `RVSearch` summary plot for HD 141004. See Figure 2.6 for plot description.

Additionally, we used Gaia DR2 to search for bound stellar companions within 10", and found no such companions. We conclude that 14 Her does not have a bound stellar companion. Therefore, we designated 14 Her c as an eccentric, long-period planet. We will continue to monitor this star with Keck/HIRES and APF, to further constrain the orbit of this planet.

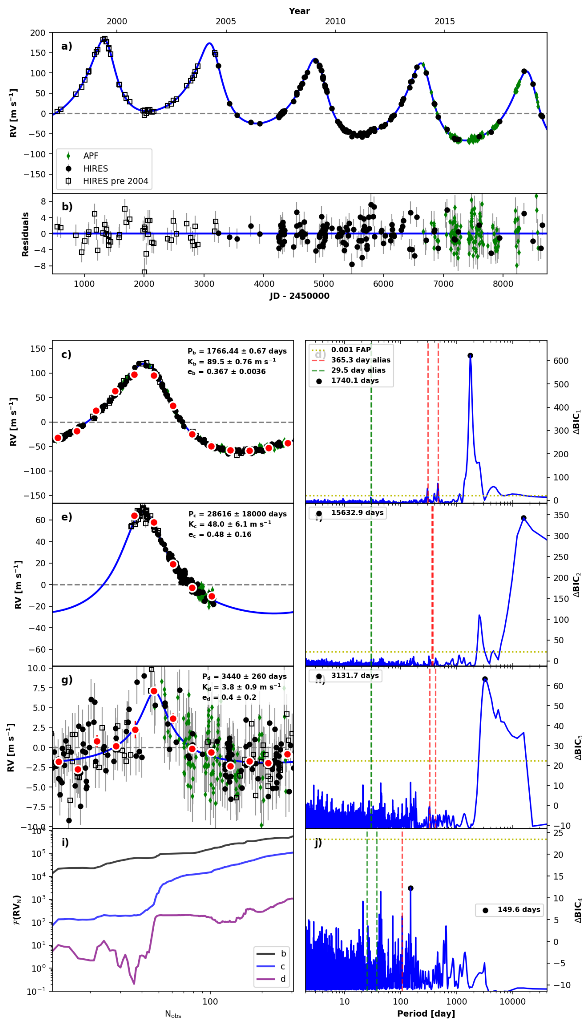

Figure 2.32: `RVSearch` summary plot for HD 145675. See Figure 2.6 for plot description.

HD 156668

HD 156668 is a K3 dwarf found at a distance of 24.4 pc. Figure 2.35 shows the `RVSearch` results for this star. This system has one known short-period super-Earth, with $M \sin i$= 4.15 $M_\oplus$and an orbital period of 4.64 d. This planet was first reported in Howard, Johnson, Marcy, Fischer, Wright, Henry, et al., 2011, which also noted a long-period ($P \approx 2.3$ yr) signal with insufficient RV observations or additional data for confirmation as a planet. The observational baseline has since increased from five years to fourteen, allowing us to confirm this long-period signal as a planet with $M \sin i$ = 0.167 M_J and an orbital period P = 2.22 yr.

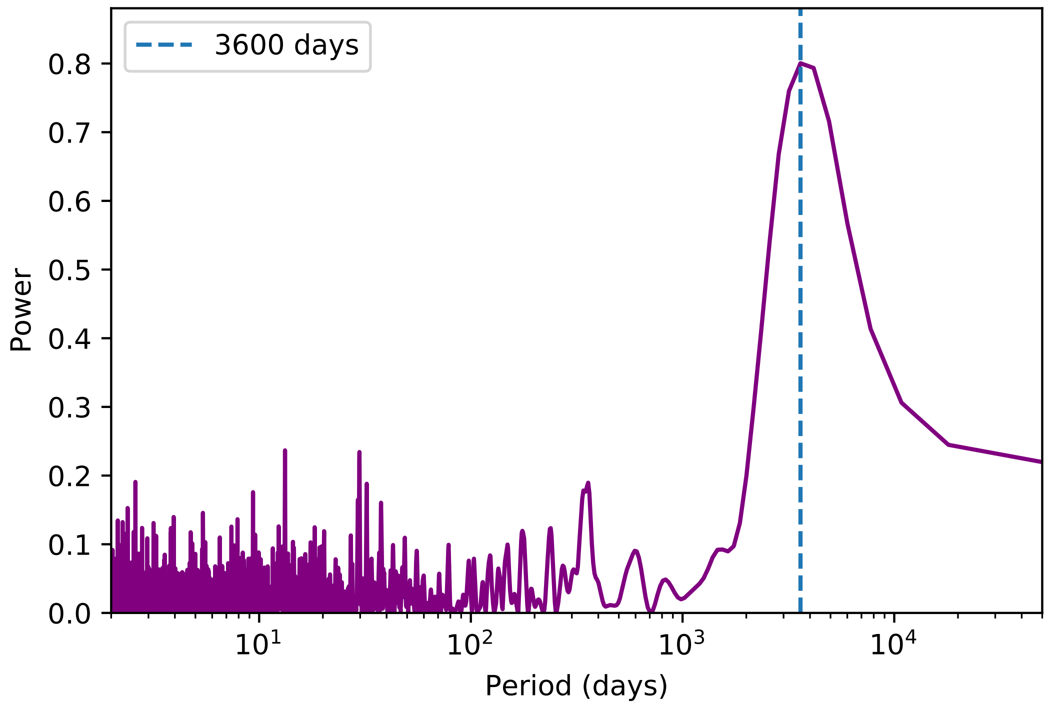

Figure 2.33: Lomb-Scargle periodogram of HIRES S-values for HD 145675 showing significant power at 3,600 days.

There is a strong periodicity in the HIRES S-value time series, peaking around 10 yr, but no significant power near the orbital period of the long-period candidate. If we do not model this activity, a one-year alias signal appears in the periodogram search (Fig. 2.35). The data does not sufficiently constrain a Keplerian fit with a 10 yr period, but we find that a linear trend models the activity well enough to remove the one-year alias from the search. We opt to include this linear trend, which we treat as a nuisance parameter.

HD 164922

HD 164922 is a G9 V dwarf located at a distance of 22.1 pc. Figure 2.36 shows the `RVSearch` results for this star. It hosts two known planets: a 0.3 M_{J} planet with an orbital period of 1207 days (Butler, Wright, et al., 2006) and a super-Earth with $M \sin i$ = 14.3 $M_{\oplus}$ and an orbital period of 75.8 days. This super-Earth was reported in Fulton, Howard, et al., 2016, which also reported residual power around 41.7 days, but did not find it significant enough to merit candidate status. With approximately two more years of HIRES and APF data, we identified the 41.7 day signal as a strong planet candidate and confirmed the 12.5-day planet reported in

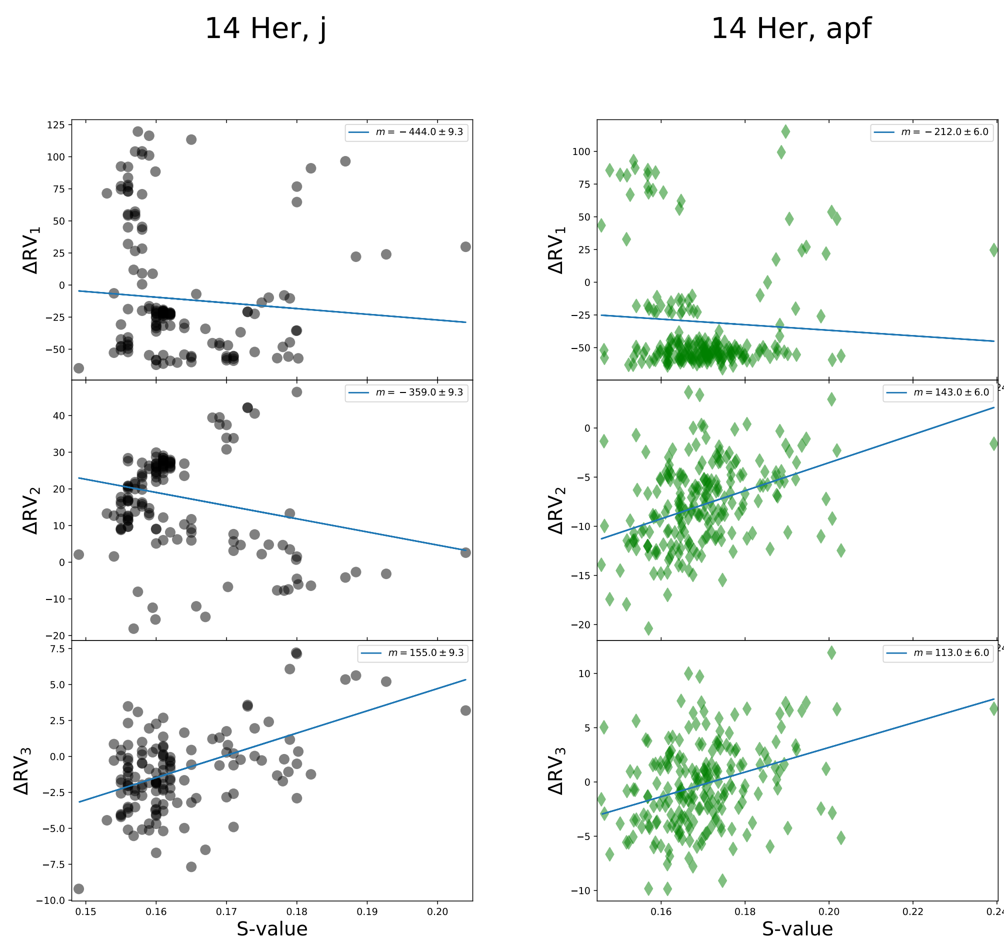

Figure 2.34: Activity vetting plots for HD 145675. For all panels, the horizontal axis shows the S-value activity metric of each observation, while the vertical axis shows corresponding RV residuals for each individual Keplerian orbit. The left-hand panels show HIRES post-upgrade observations, while the right-hand panels show APF observations. Each row shows RVs with the model residuals of one Keplerian model, with the other Keplerian models subtracted from the data. The blue lines show linear correlations between these residuals and the corresponding S-values. In the HIRES and APF data, we measured $> 3\sigma$ correlations for the third Keplerian signal. The APF and HIRES linear correlations are within 3σ of each other, implying that this signal is caused by stellar activity. We find correlations between the residuals and S-values for the second signal as well, but they are significantly different for HIRES and APF. Since the period of this signal is much greater than the APF baseline of this star, we discount this second correlation as caused by the limited baseline of the data with respect to the signal.

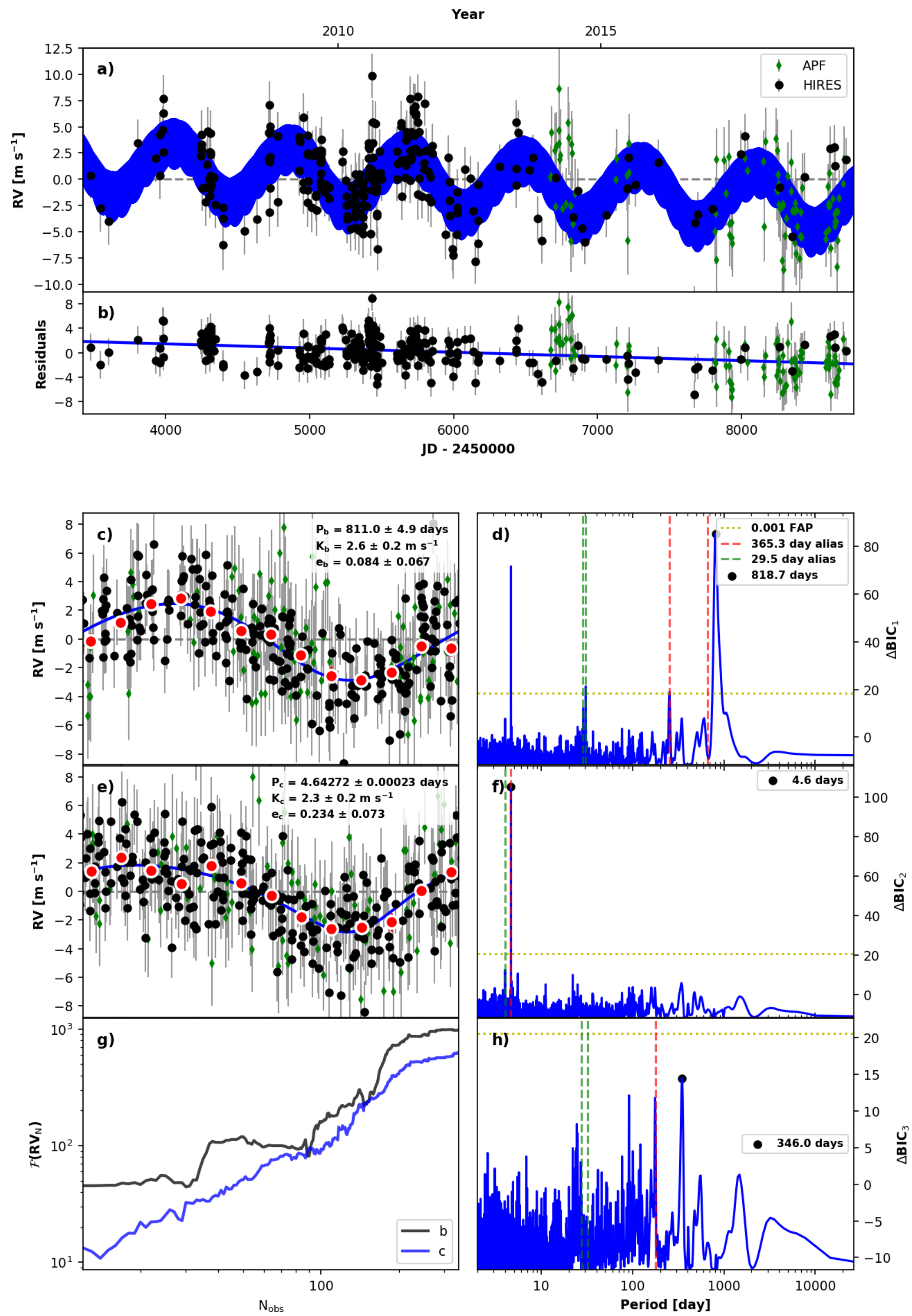

Figure 2.35: `RVSearch` summary plot for HD 156668. See Figure 2.6 for plot description.

Benatti et al., 2020. Both planets are of sub-Neptune mass, and have eccentricity posteriors that are consistent with circular orbits. The 41.7 day planet has $M \sin i = 10.7 \pm 1.0\ M_\oplus$ and a semi-major axis $a = 0.2294 \pm 0.0031$ AU. The 12.5 day planet has $M \sin i = 4.63 \pm 0.70\ M_\oplus$ and a semi-major axis $a = 0.1024 \pm 0.0014$ AU. Table 1 reports all planet parameters.

To validate these candidates, we searched for periodicity in both S-value activity metrics and APT photometry. We found no evidence for stellar rotation in S-values, but estimated a stellar rotation period of 62.1 days from our APT photometry. Figure 2.37 shows periodograms and a phase-folded curve from this APT analysis, and Figure 2.38 shows equivalent analysis for HIRES S-values. The 1-year alias of 62.1 days is 75.8 days, but the 75.8 day planet detection is high-amplitude and clean, without an additional peak near 62 days in any of the `RVSearch` periodograms. Therefore, within the limits of our activity metrics and vetting process, we ruled out stellar rotation as a cause of the 41.7-day signal.

Benatti et al., 2020 used multiple HARPS-N spectral activity indicators to estimate a stellar rotation period of 41.6 days, and notes that this rotation period is to be expected from empirical activity-rotation relationships. Therefore, they determined that the strong 42-day signal present in their HARPS RVs is caused by rotation. However, we find no evidence of significant 42-day periodicity in our analysis of spectral activity indicators or APT photometry, as seen in Figures 2.37 and 2.38, and both datasets reflect significant periodicity near 60 days. Since our RV detection of this planet candidate is clean and does not conflict with our activity analysis, we chose to include this signal in our catalog as a planet candidate, to be confirmed or refuted by independent analysis.

HD 168009

HD 168009 is a G1 dwarf found at a distance of 23.3 pc. Figure 2.39 shows the RVSearch results for this star. `RVSearch` discovered a super-Earth candidate at an orbital period of 15.5 days, with $M \sin i = 10.3\ \pm 1.1\ M_\oplus$. Table 1 reports current estimates of all planet parameters.

`RVSearch` also recovered a highly eccentric 1-year signal, but this signal correlates with APF systematics. Therefore, we labeled it as a false positive.

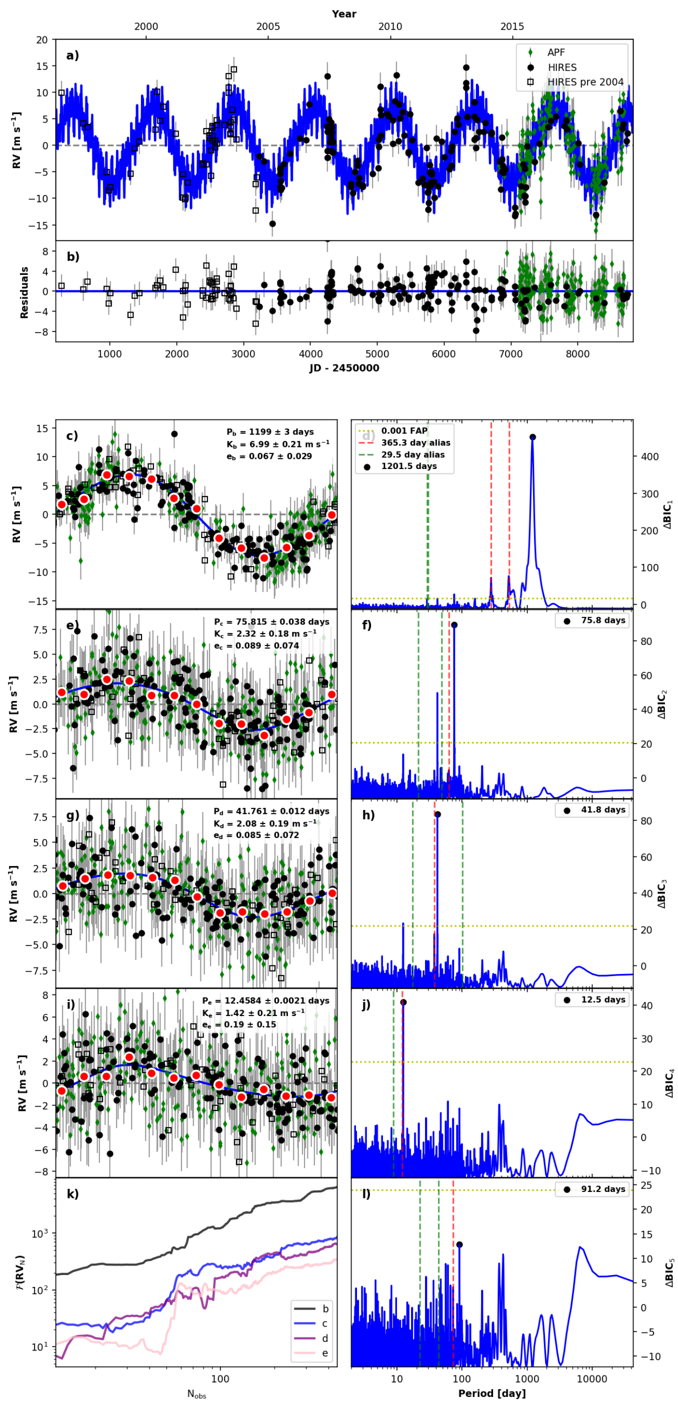

Figure 2.36: `RVSearch` summary plot for HD 164922. See Figure 2.6 for plot description.

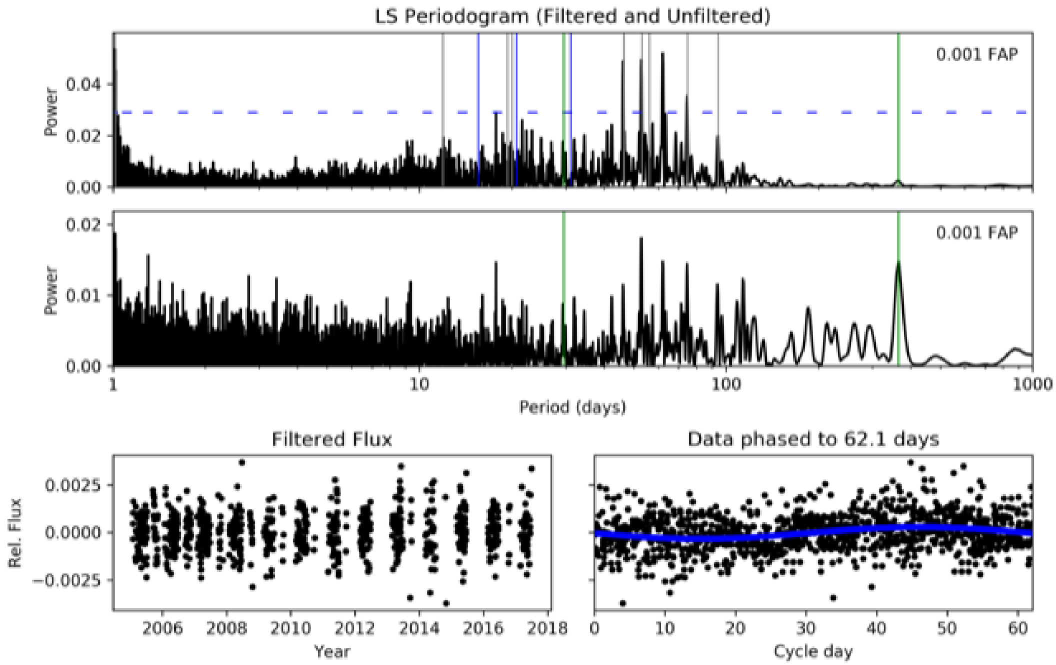

Figure 2.37: Visualization of APT photometry analysis for HD 164922. The top panel shows a Lomb-Scargle periodogram of the photometry, with a moving-average filter to reduce alias issues. The middle panel shows an unfiltered periodogram.

HD 213472

HD 213472 is a G5 dwarf located at a distance of 64.6 pc. Figure 2.40 shows the `RVSearch` results for this star. There is an approximately eleven-year gap in RV observations of this star. The first post-upgrade HIRES observation was measured in 2005, shortly after the last pre-upgrade observation, and the second post-upgrade observation was measured in 2016. The 40 m s^{-1} difference between these two observations prompted the CPS team to begin observing HD 213472 regularly. Together with observations since 2016, and the thirteen pre-upgrade HIRES measurements, the data are consistent with a long-period, eccentric, planetary companion. Our periodogram search detects such a long-period signal. Due to the sparseness of the data, traditional MCMC methods fail to return a well-sampled model posterior. We used the rejection sampling algorithm `TheJoker` (Price-Whelan et al., 2017) to estimate the posterior, and found it to be unimodal. This model is consistent with a very long-period gas giant, with $M \sin i = 3.48^{+1.10}_{-0.59}\ M_{\mathrm{J}}$ orbital period $P = 46^{+33}_{-13}$ yr, semi-major axis $a = 13.0^{+5.7}_{-2.6}$ AU, and eccentricity of $e = 0.53^{+0.12}_{-0.09}$. Table 1 reports all planet parameters. Figure 2.41 shows the orbital parameter posteriors generated by TheJoker.

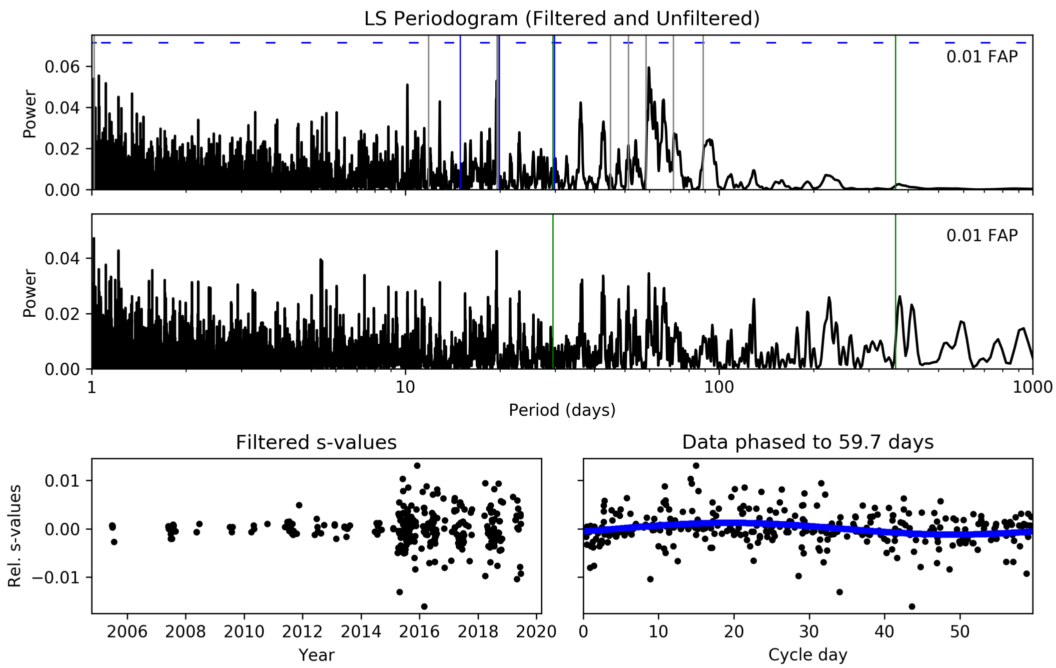

Figure 2.38: Visualization of HIRES S-value analysis for HD 164922. The top panel shows a Lomb-Scargle periodogram of the S-values, with a moving-average filter to reduce alias issues. The middle panel shows an unfiltered periodogram.

To investigate the possibility of a stellar or substellar companion, we compared this Keplerian model to a simple linear trend, by computing the ΔBIC between the two max-likelihood models. The Keplerian model is significantly preferred with ΔBIC = 23.7. Additionally, we used Gaia to search for bound companions within 10", and found no such companions. Therefore, we inferred that HD 213472 b is either a planet or low-mass substellar companion, and not a wide-orbit stellar companion.

Figure 2.42 shows a sample of possible orbits for HD 213472 b, drawn from our rejection sampling posteriors and projected over the next decade. More HIRES observations will further constrain this object's mass and orbital parameters.

Stellar, Planet, Trend, and Data Catalogs

We have compiled several tables of catalog information: stellar properties, detected planets, long-term RV trends, false positives, and RV datasets. These tables are available at `github.com/leerosenthalj/CLSI`.

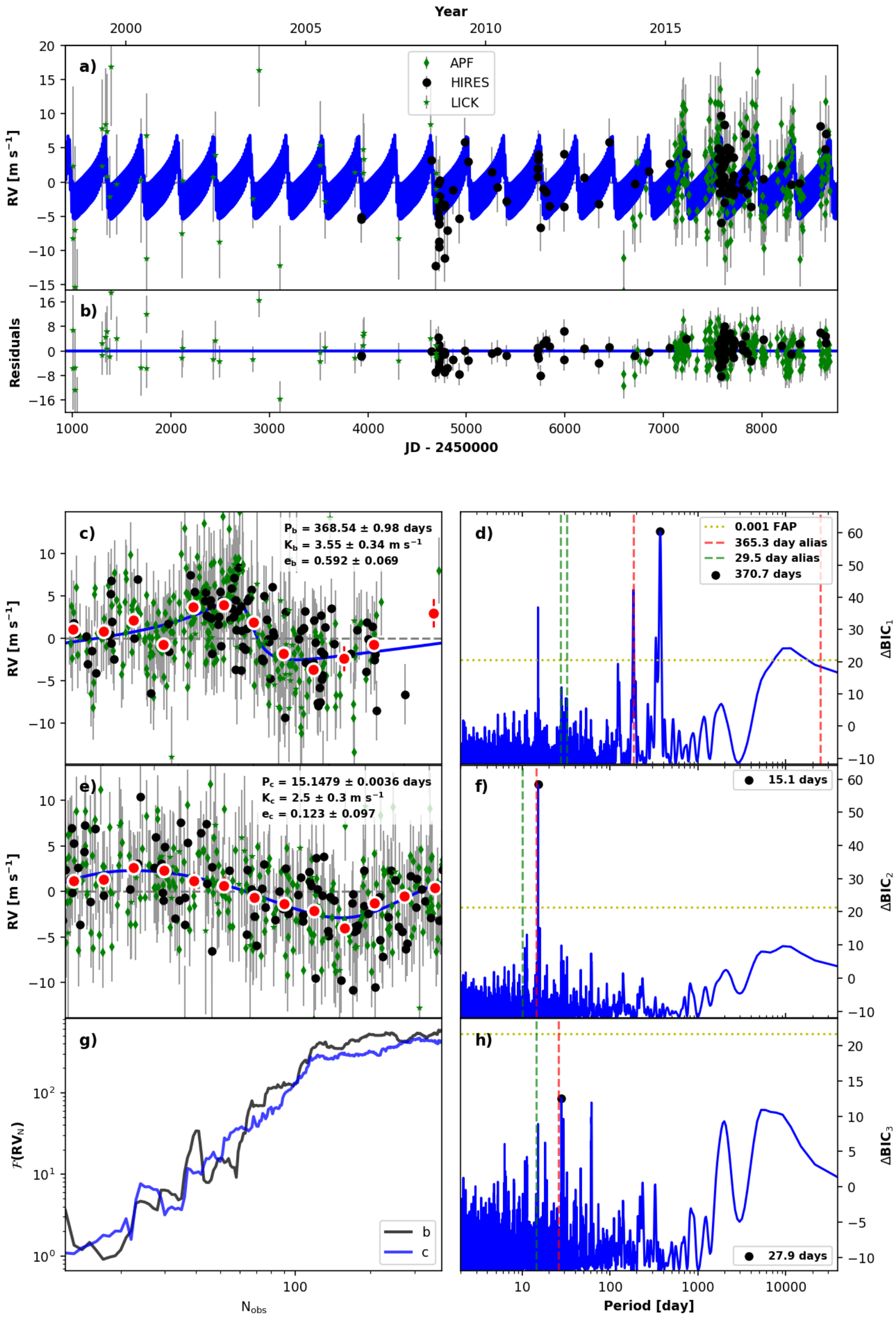

Figure 2.39: RVSearch summary plot for HD 168009. See Figure 2.6 for plot description.

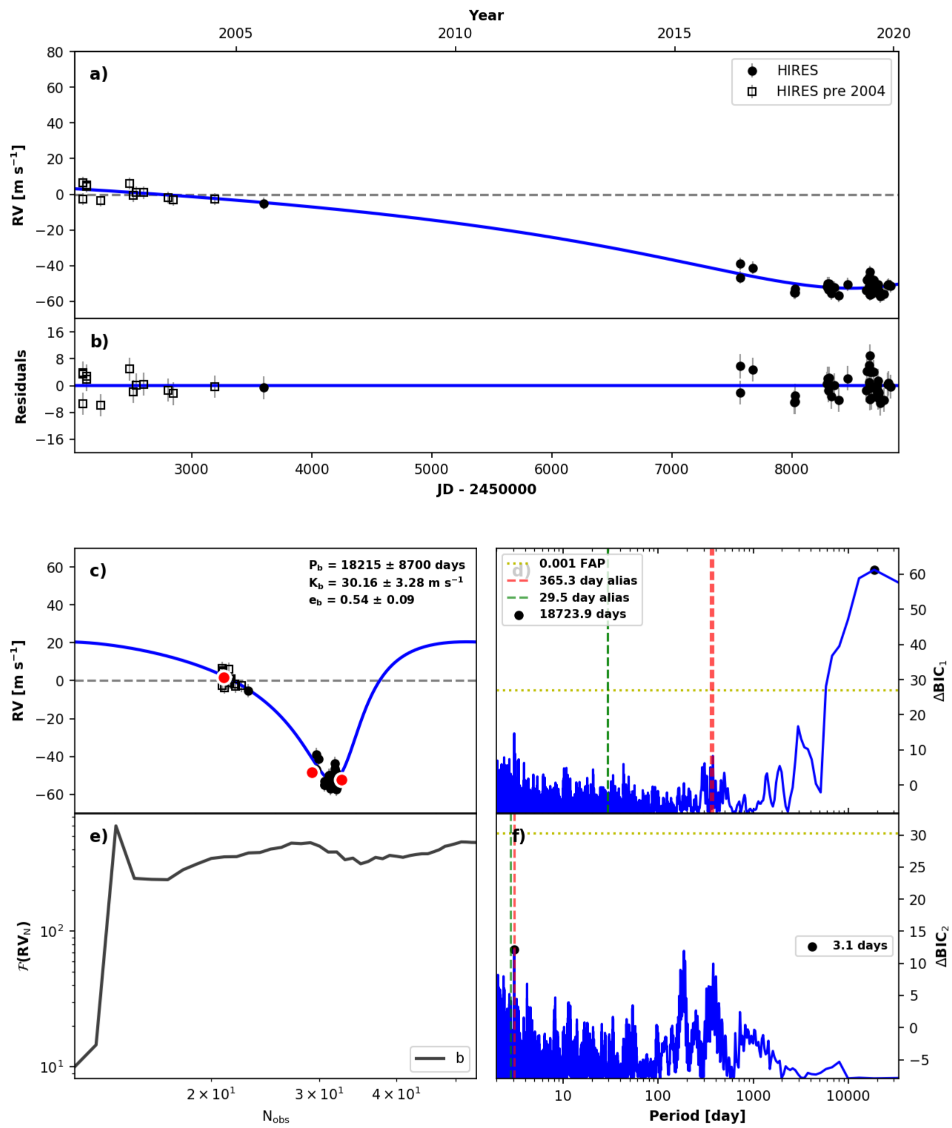

Figure 2.40: `RVSearch` summary plot for HD 213472. See Figure 2.6 for plot description.

False Positives

Through our false positive vetting process, we found evidence for several false positives that correspond to reported planets in the literature, or to stars that have been discussed extensively in the literature. We elaborate on each of these cases in the subsections below.

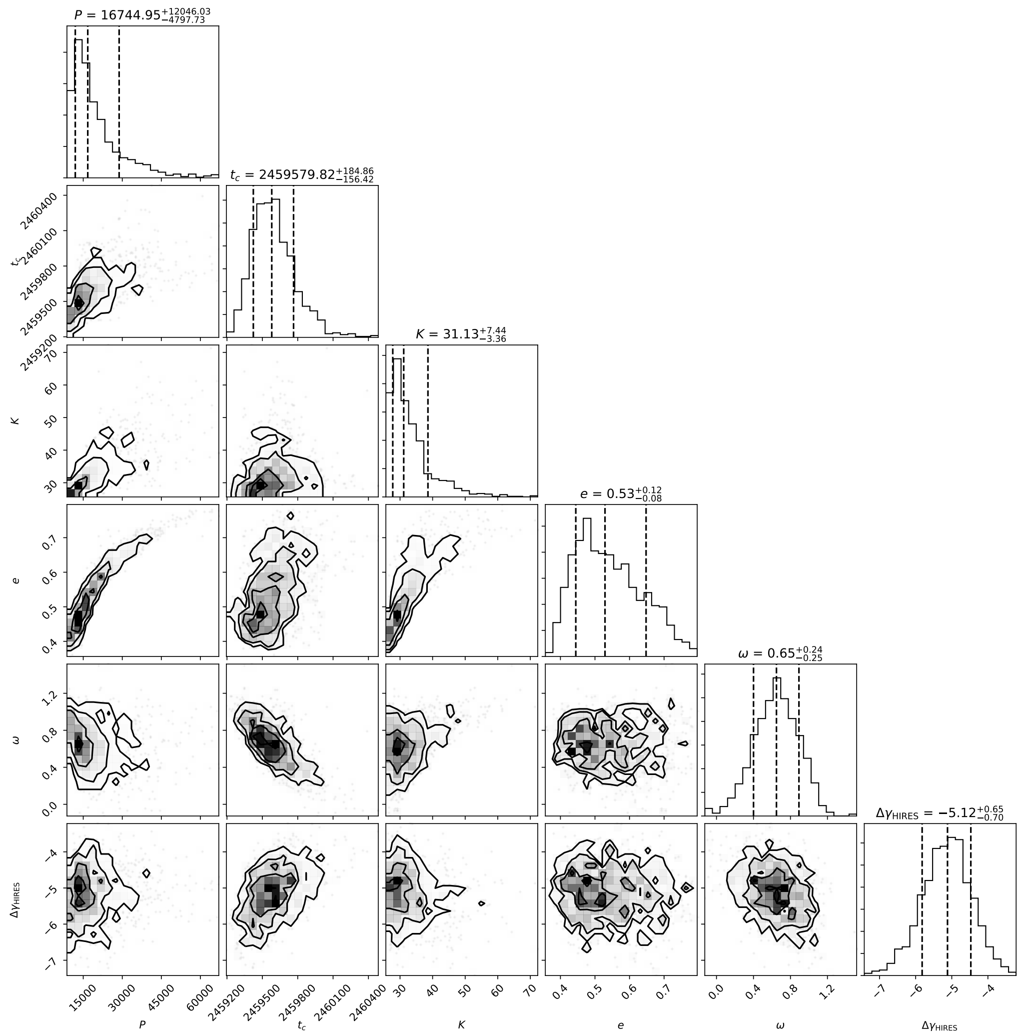

Figure 2.41: Rejection sampling posterior for HD 213472 b orbital parameters. $\Delta\gamma$ is the relative linear offset between different instrumental datasets, in this case pre-upgrade and post-upgrade HIRES.

HD 115617

Vogt, Wittenmyer, et al., 2010 reported three planets orbiting this star, with periods of 4.2, 38, and 124 days. `RVSearch` recovered signals at all three periods. However, the 124 day signal (1/3rd of a year) has a strong harmonic at 1/4th of a year, and there is significant residual power at roughly one year, as seen in panels h and j of Figure 2.43. We investigated this candidate by computing periodograms for the 12 HIRES PSF parameters computed for each RV measurement, and found periodicity at 1 yr and harmonics of 1 yr for several parameters, as seen in Figure 2.44. Additionally, several of these PSF parameters correlate strongly with the corresponding RVs,

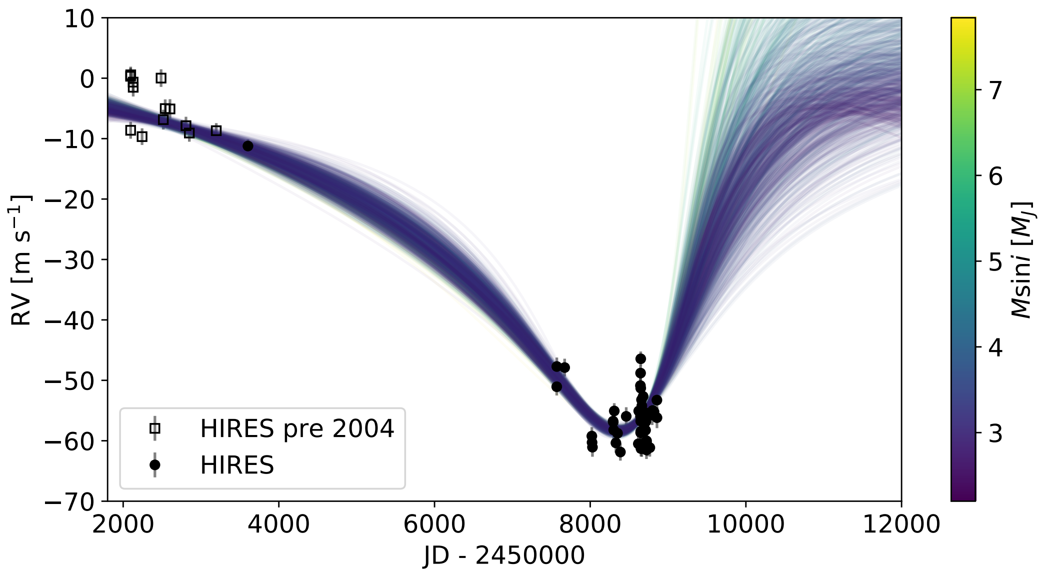

Figure 2.42: Possible orbits for HD 213472 b. RV curves are drawn from the rejection sampling posterior generated with `TheJoker`. The color of each orbit drawn from the posterior scales with $M \sin i$.

after subtracting the RV models of the two inner planets, as seen in Figure 2.45. Therefore, we designated the 124 day signal as a yearly systematic.

HD 154345

Here, we confirm the planetary status of the planet claim for HD 154345. Wright, Marcy, Butler, et al., 2008 announced the detection of a true Jupiter analog, with $M \sin i = 0.95\ M_{\rm J}$ and an orbital period of 9.2 yr, corresponding to an orbital separation of 4.2 AU. This chapter also presented strong evidence for a stellar magnetic activity cycle with a periodic timescale of roughly nine years. As the CPS group continued to observe HD 154345 over the next few years, the planet candidate's RV signature and the corresponding S-values appeared to be strongly in phase, and Wright, 2016 noted that the candidate may be a false positive. However, in the twelve years since HD 154345 b was initially reported, HIRES RV measurements and activity metrics have drifted from being completely in phase to being completely out of phase, as seen in Figure 2.46, and therefore are not linearly correlated. This strongly implies that this Jupiter analog candidate cannot be attributed to stellar activity, and that this candidate should be cemented as a confirmed planet. `RVSearch` detects two signals in our HD 154345 dataset, both close to 9 yr, as seen in Figure 2.47. We attribute the circular orbit with a greater RV amplitude to HD 154345 b, and the weak, eccentric signal to stellar activity.

HD 26965

Ma et al., 2018 reported a 42.4 day super-Earth orbiting the nearby star HD 26965, using datasets taken by multiple spectrographs, including HIRES. We detected significant periodicity at 42 days in the HIRES S-value measurements as seen in Figure 2.48, and determined that 42 days is the likely stellar rotation period of HD 26965. There is also evidence of a long-period magnetic activity cycle, as seen in the juxtaposition of S-values and RVs in Figure 2.49.

HD 34445

Howard, Johnson, Marcy, Fischer, Wright, Bernat, et al., 2010 reported a giant planet orbiting this star at a period of 1049 days. Vogt, Butler, Burt, et al., 2017 reported five small planets, claiming evidence in LCES-derived HIRES radial velocities. `RVSearch` detected the giant planet and three of the five small planet claims, as seen in the summary plot shown in Figure 2.50. The longest-period candidate among the five, not modeled as a Keplerian here, clearly correlates with HIRES S-values; we model this signal with a linear trend, for simplicity. Figure 2.51 juxtaposes the HIRES S-values and corresponding RVs, minus the Keplerian signal of the system's giant planet. As for the three other periodic signals that we detect, two are likely HIRES systematics and one is likely stellar rotation We detected significant periodicity at 52 days in the HIRES S-value measurements as seen in Figure 2.52, and determined that 52 days is the likely stellar rotation period of HD 34445. This places our weak detection of the 49 day claimed planet candidate under suspicion, and we have labeled it as a false positive in our catalog. There is also evidence of semiannual HIRES systematics, as seen in Figure 2.54, which shows the correlation between HIRES RVs minus the giant planet signature and PSF parameters, and in Figure 2.53, which shows periodograms of each PSF parameter time series. Multiple PSF parameters correlate ($|R| > 0.15$) with the RV residuals, and multiple parameters show periodicity around one-third and one-fourth of a year. The two claimed planets at 118 and 215 days are close to one-third and one-half of a year, respectively, and show weak and equal-strength signatures in their `RVSearch` periodograms, as seen in Figure 2.50. Therefore, we have labeled these signals as false positives in our catalog.

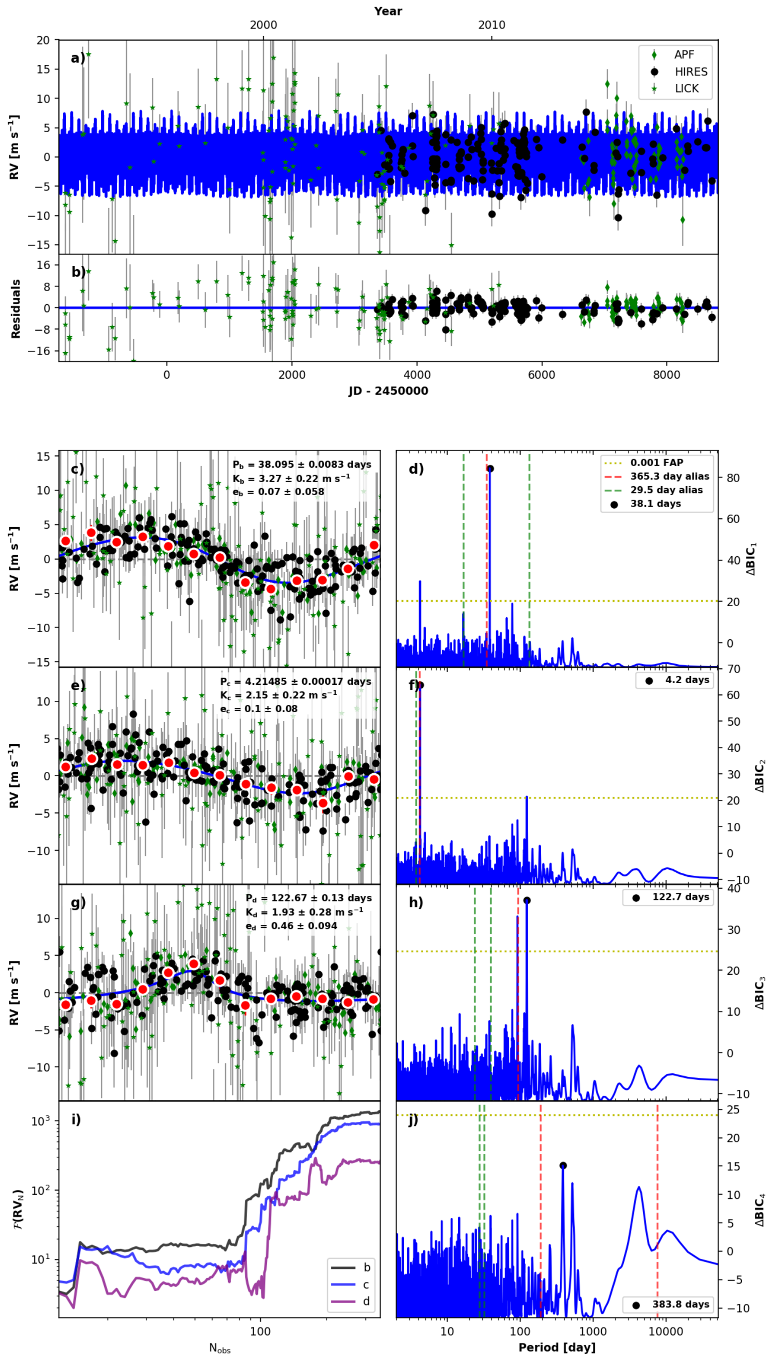

Figure 2.43: `RVSearch` summary plot for HD 115617. See Figure 2.6 for plot description. Note the nearly equivalent-height peaks at 1/3 and 1/4 year in panel h, corresponding to the 124 day reported planet. Panel j shows that there is residual power at 1 year after subtracting the 122 day signal, suggesting the presence of yearly systematic noise in the data.

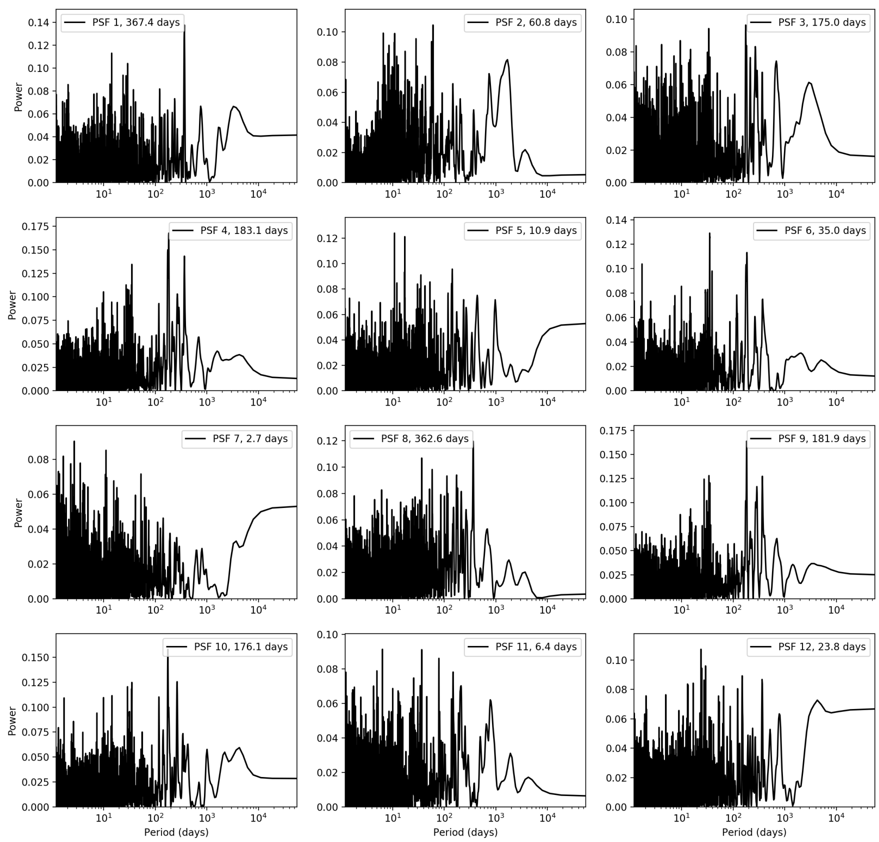

Figure 2.44: PSF Lomb–Scargle periodograms for HD 115617. Each panel corresponds to a Doppler code PSF fitting parameter.

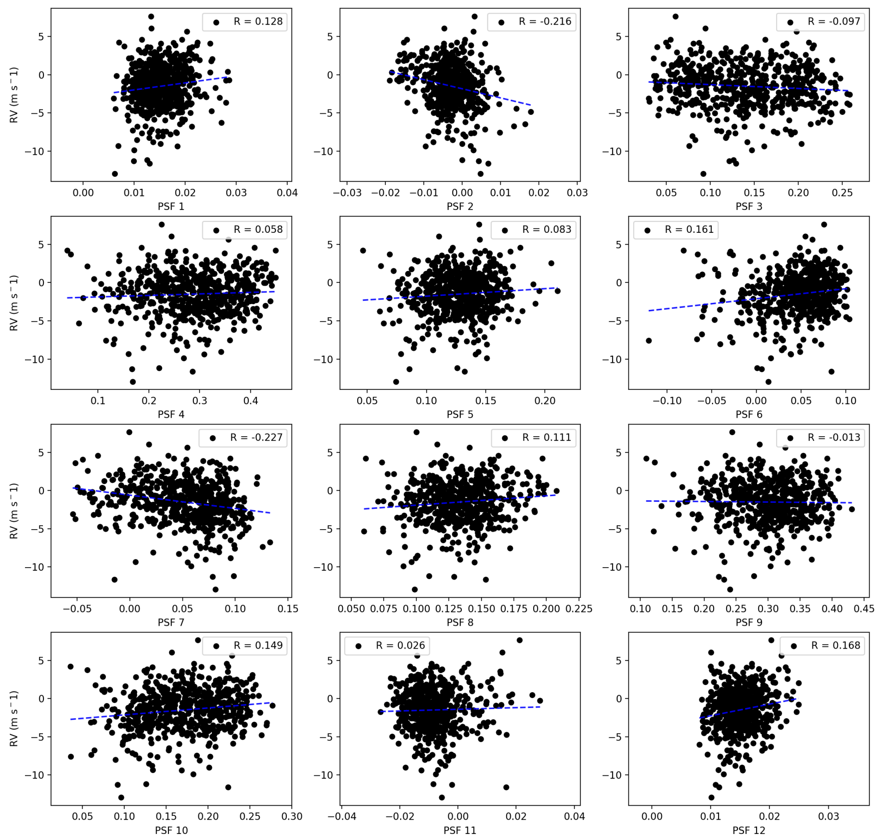

Figure 2.45: PSF correlation plots for the candidate HD 115617 d. Each panel corresponds to a Doppler code PSF fitting parameter, with PSF value on the x-axis and RV without the signatures of the inner two planets on the y-axis. Dashed blue lines are least-squares linear fits. R is the Pearson correlation value; multiple PSF parameters have $|R| > 0.15$.

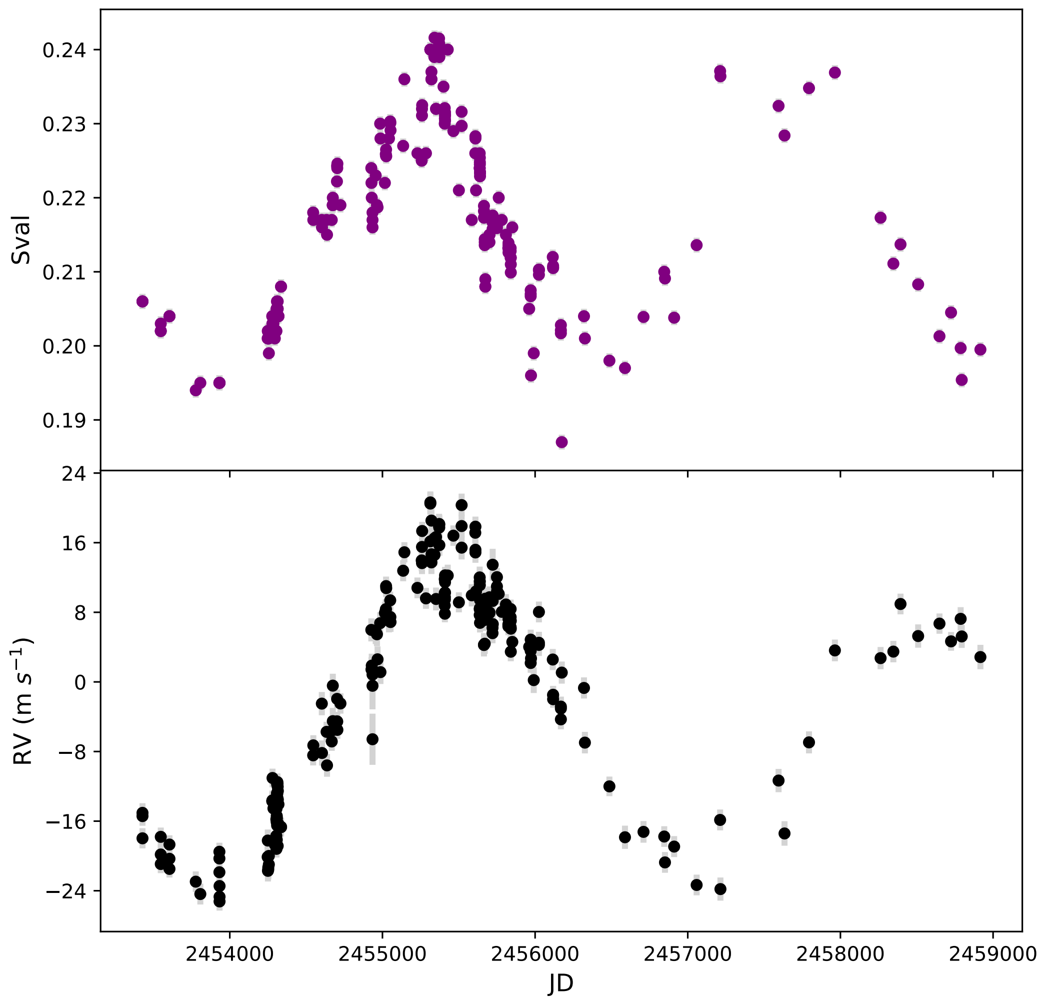

Figure 2.46: HIRES post-upgrade RV and S-value activity timeseries for HD 154345. Note that the two datasets share minima and appear to be in phase when post-upgrade observations began, but have drifted completely out of phase over the following 23 years.

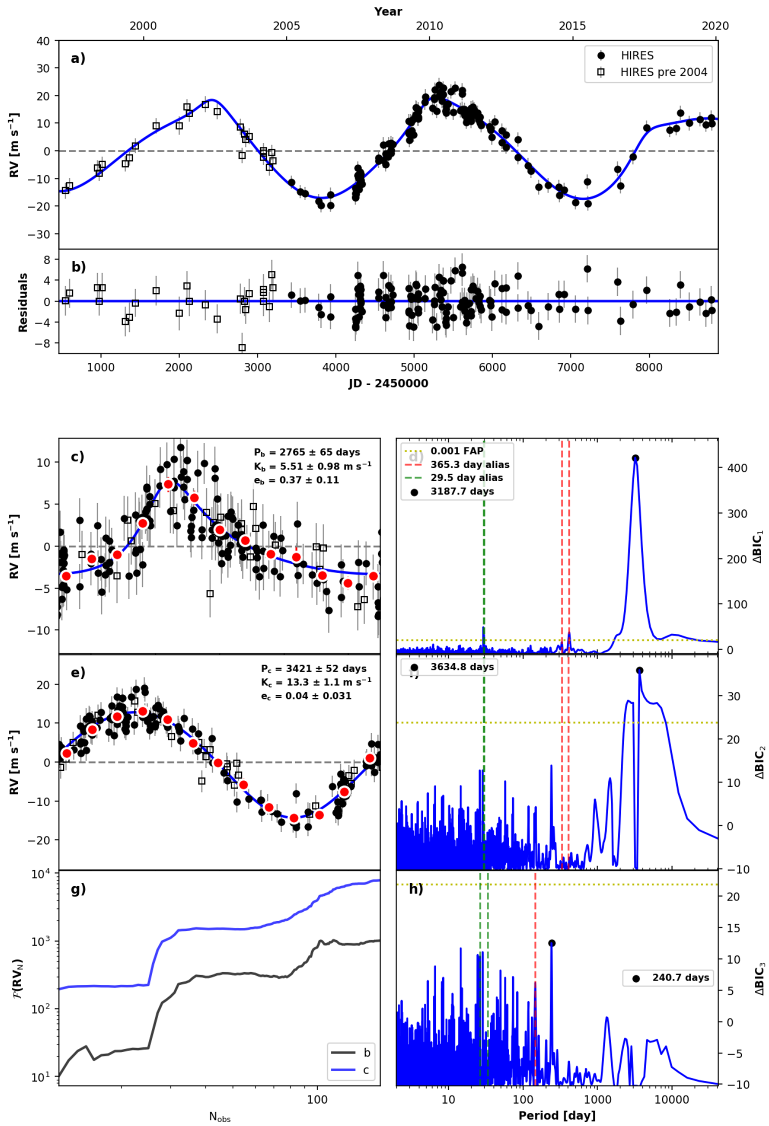

Figure 2.47: `RVSearch` summary plot for HD 154345; see Figure 2.6 for description. `RVSearch` first recovered a strong signal at 9 years, but then recovered additional power at a similar period due to stellar activity. The final orbit fit switched the two models, so that panels e) and d) show the planetary signal, while panels c) and f) show the stellar activity signal.

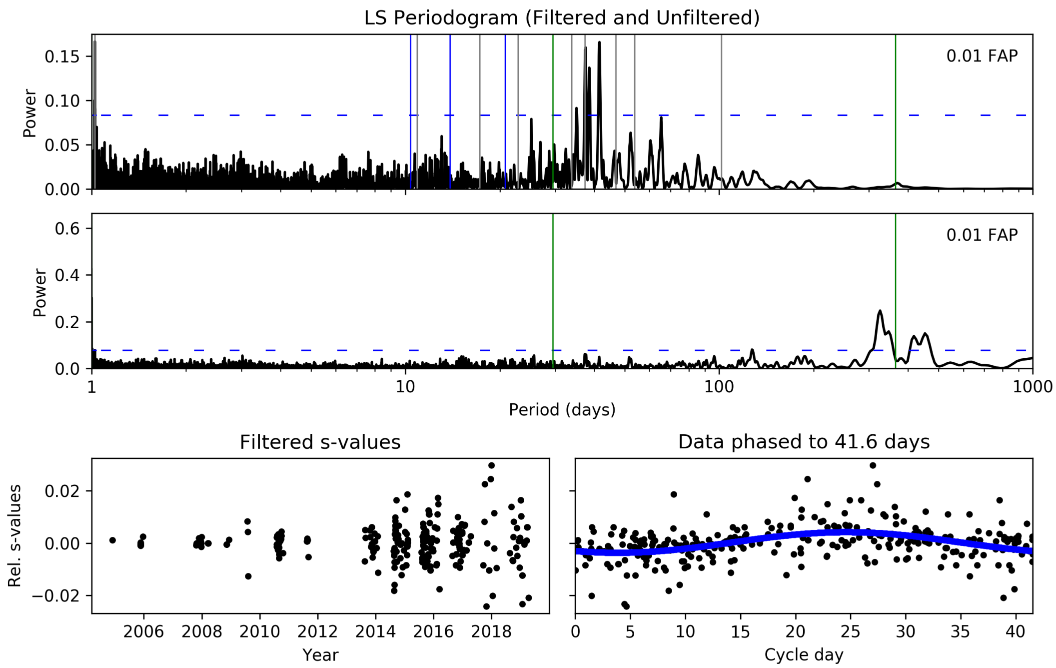

Figure 2.48: Stellar rotation analysis of HIRES S-values for HD 26965. The top panel shows a Lomb–Scargle periodogram of the S-values after we applied a high-pass filter to them, to remove the impact of the long-period magnetic activity cycle. The middle panel shows a periodogram of the raw S-values. The top panel shows significant periodicity near 40 days, with a maximum at 41.6 days. The bottom-left panel shows the filtered S-values, while the bottom-right panel shows the filtered S-values phased to 41.6 days; there appears to be a coherent signal at this period, implying stellar rotation with this period.

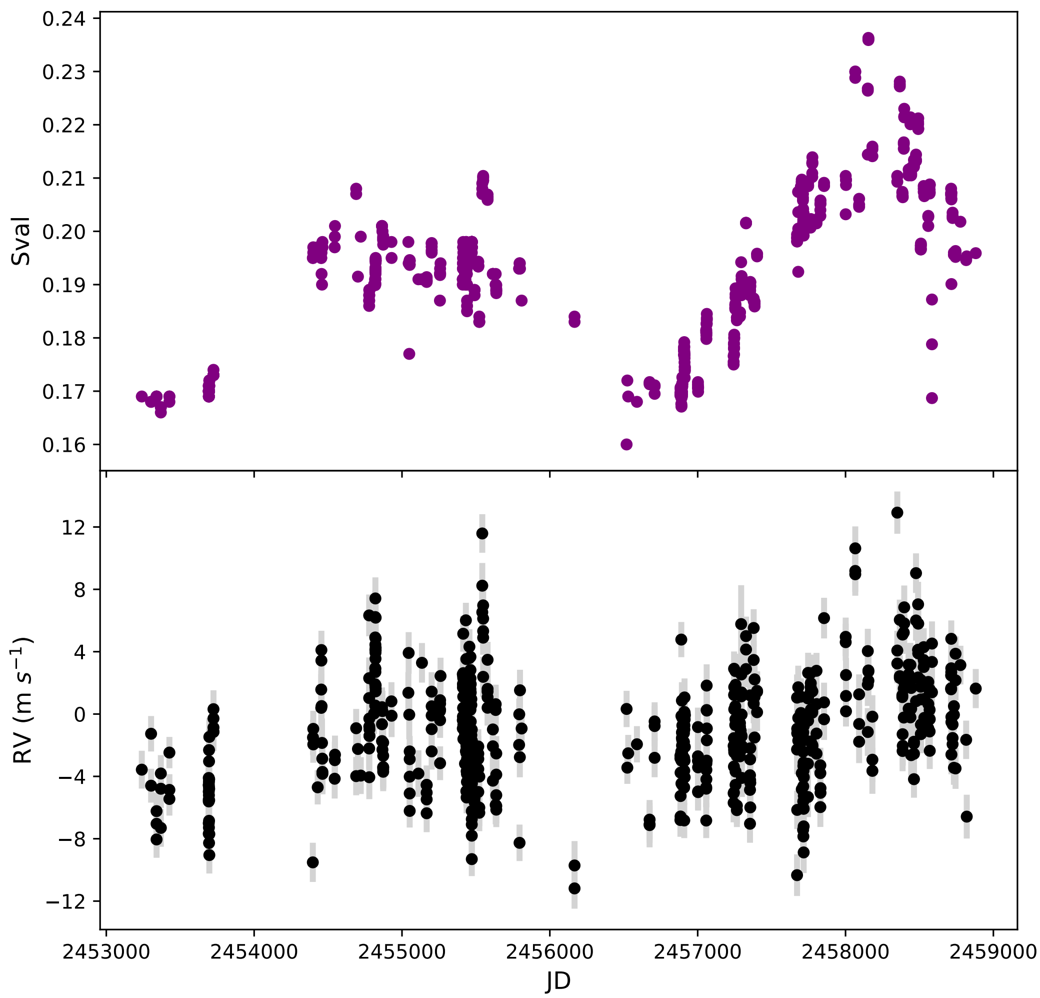

Figure 2.49: HIRES post-upgrade S-values and RVs for HD 26965. The two datasets both have long-period power and are in phase with each other.

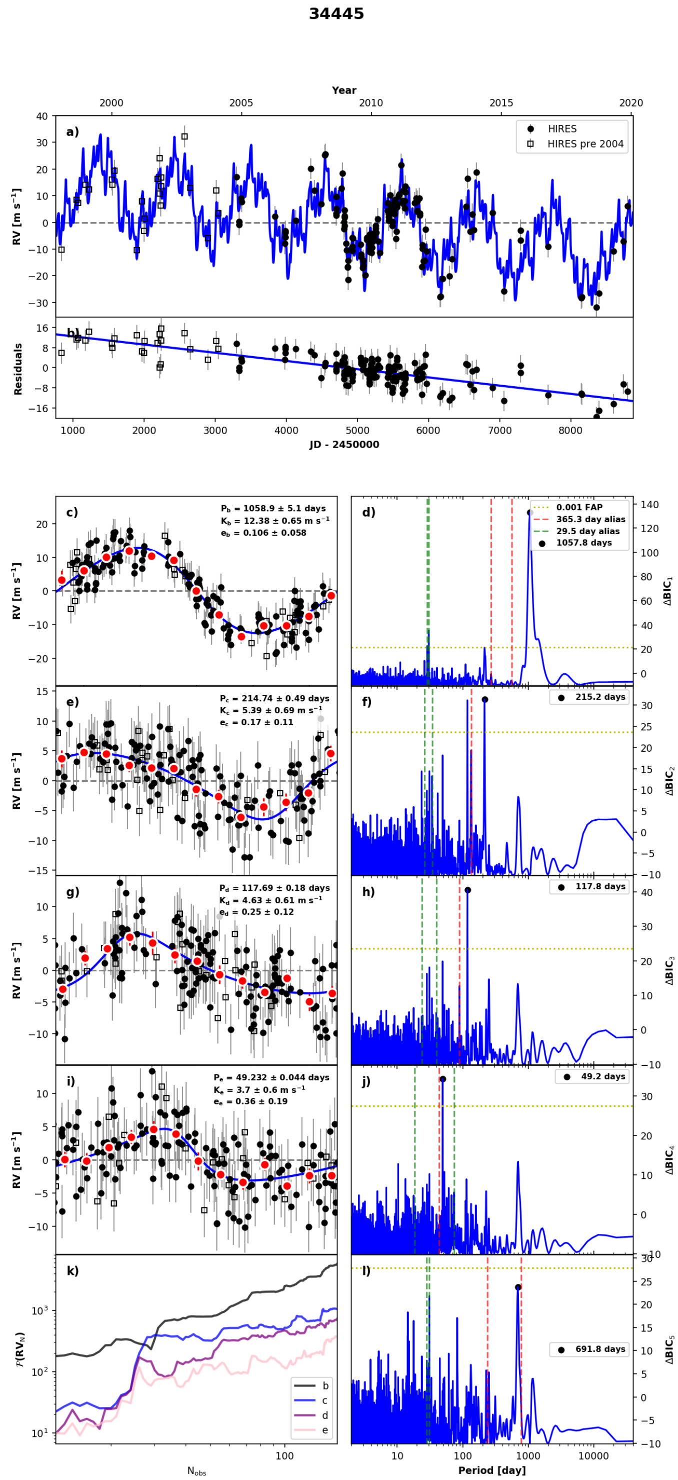

Figure 2.50: `RVSearch` summary plot for HD 34445; see Figure 2.6 for description. `RVSearch` first recovered the known giant planet, then a series of what are likely spurious signals caused by yearly aliasing or stellar activity.

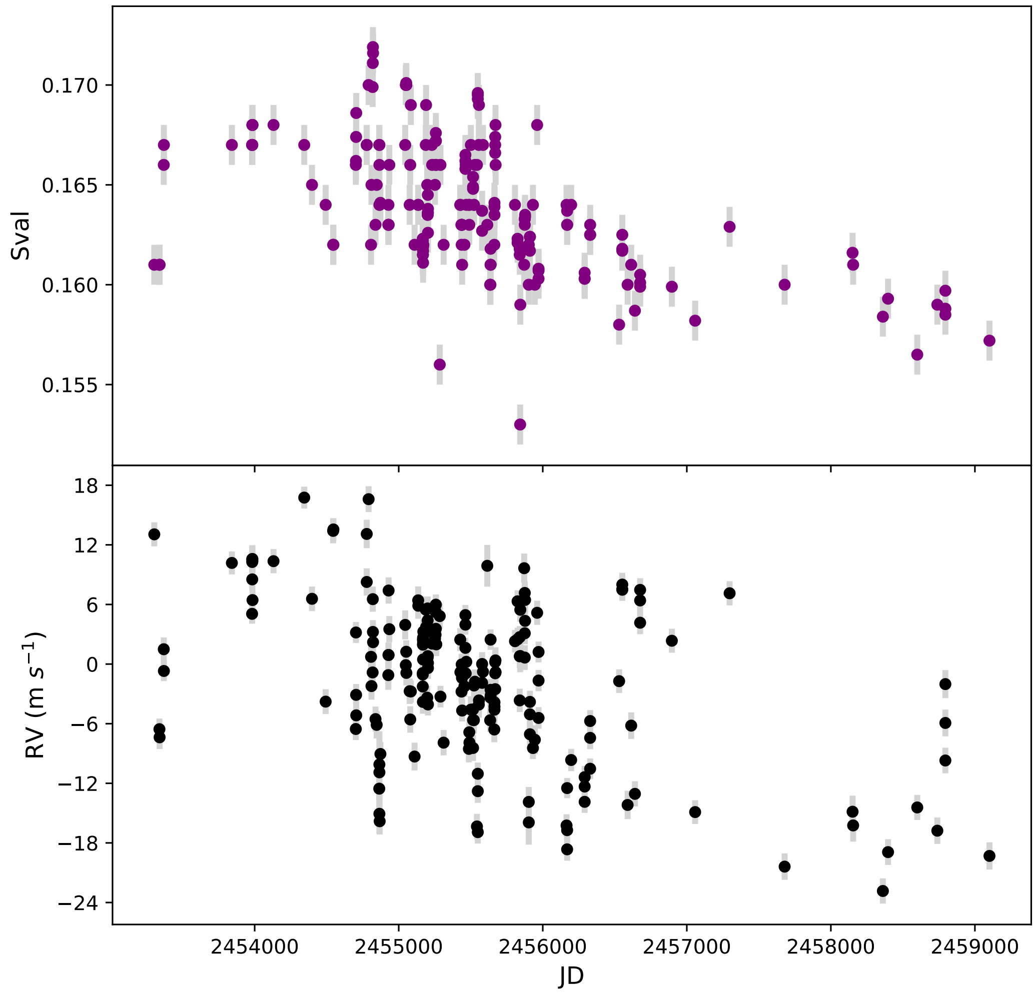

Figure 2.51: HIRES post-upgrade RV and S-value activity timeseries for HD 34445, with the giant planet RV model subtracted. Note that these two datasets share a negative long-term trend, which we believe accounts for the claimed 5,700-day planet in the system.

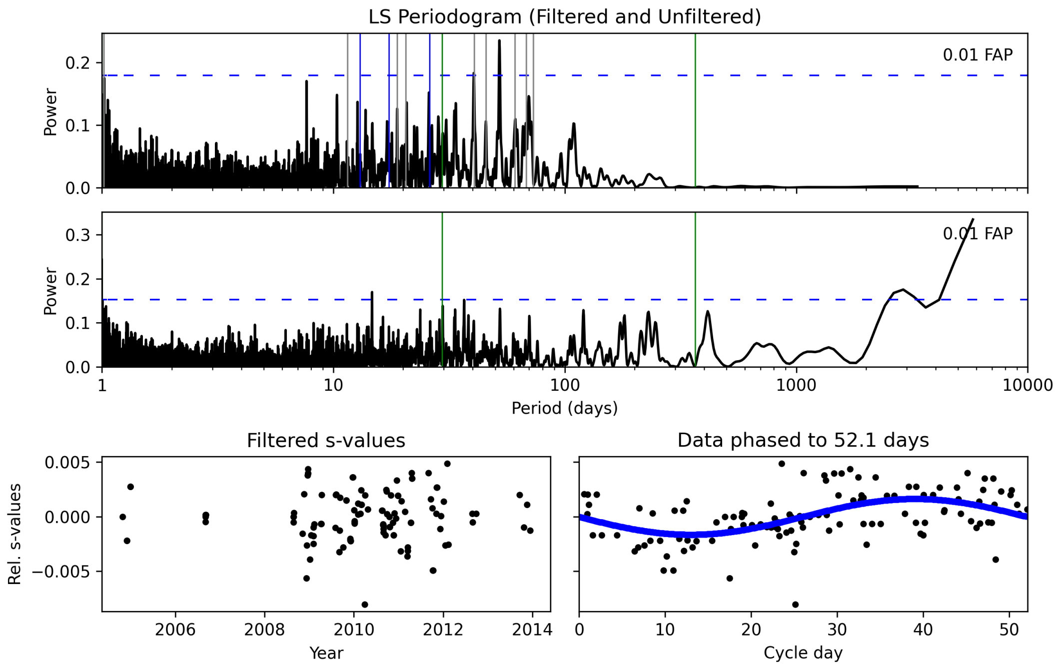

Figure 2.52: Stellar rotation analysis of HIRES S-values for HD 34445. The top panel shows a Lomb–Scargle periodogram of the S-values after we applied a high-pass filter to them, to remove the impact of the long-period magnetic activity cycle. The middle panel shows a periodogram of the raw S-values. The top panel shows significant periodicity around 52.1 days. The bottom-left panel shows the filtered S-values, while the bottom-right panel shows the filtered S-values phased to 52.1 days; there appears to be a coherent signal at this period, implying stellar rotation with this period. This led us to label the 49 day claimed planet as a false positive, since there is insufficient evidence to distinguish it from stellar rotation.

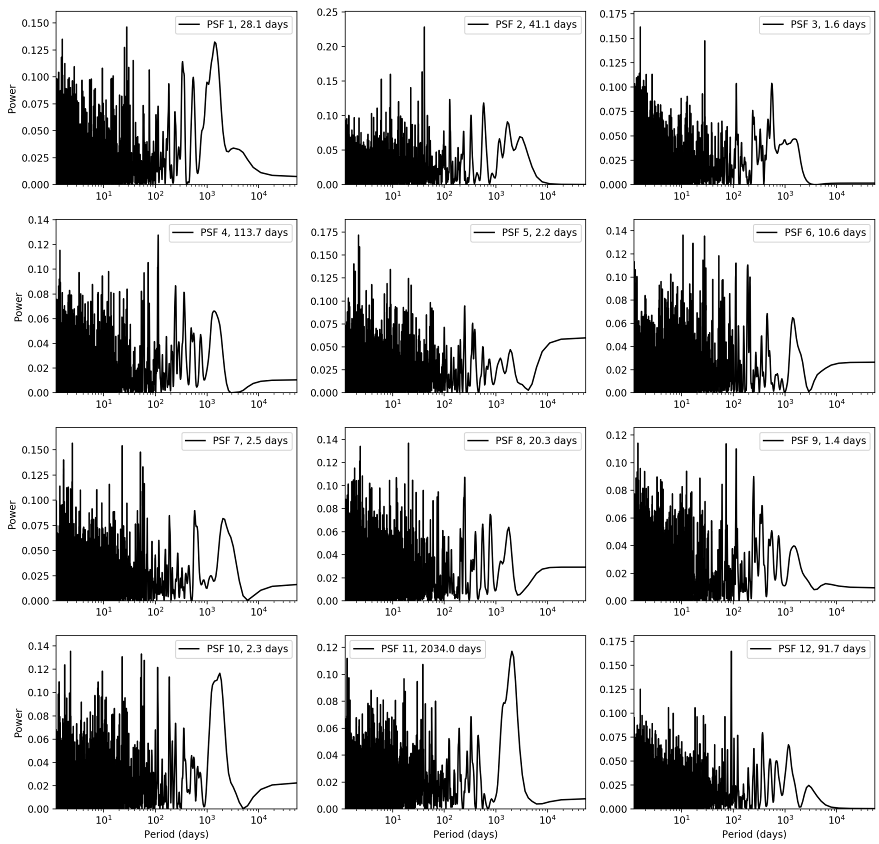

Figure 2.53: PSF Lomb–Scargle periodograms for HD 34445. Each panel corresponds to a Doppler code PSF fitting parameter.

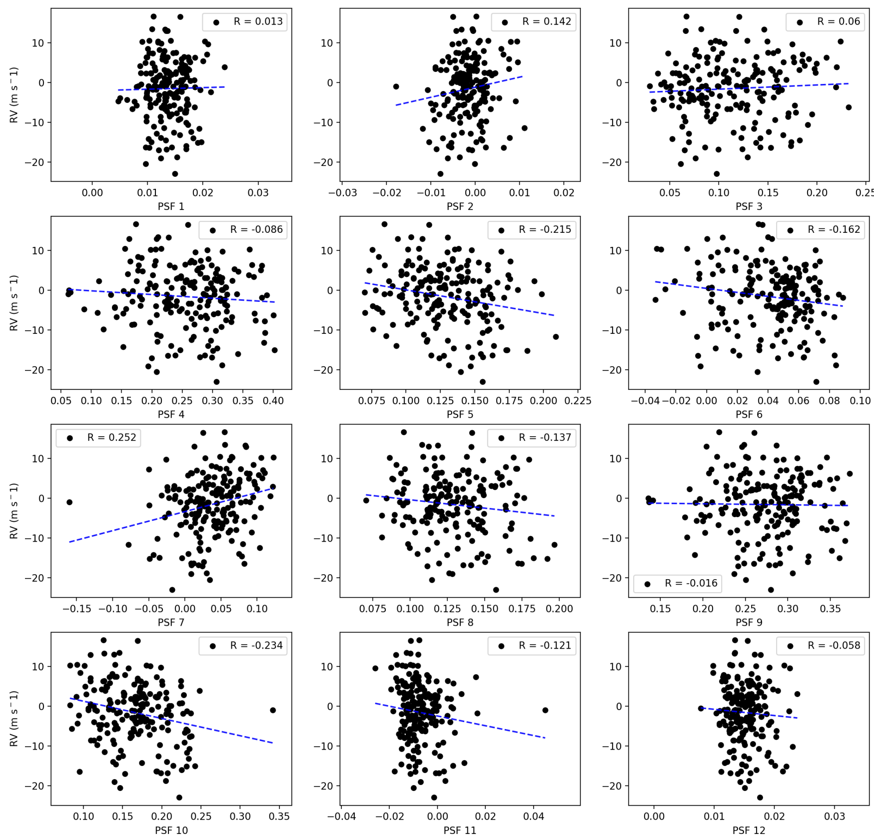

Figure 2.54: PSF correlation plots for HD 34445, without the RV signature of the star's giant planet. Each panel corresponds to a Doppler code PSF fitting parameter, with PSF value on the x-axis and RV without the giant planet signature on the y-axis. Dashed blue lines are least-squares linear fits. R is the Pearson correlation value; multiple PSF parameters have $|R| > 0.15$.

Chapter 3

FROST GIANTS: THE OCCURRENCE OF JOVIAN PLANETS BEYOND THE ICE LINE

Fulton, B. J. et al. (July 2021). “California Legacy Survey. II. Occurrence of Giant Planets beyond the Ice Line”. In: *ApJS* 255.1, 14, p. 14. DOI: 10.3847/1538-4365/abfcc1. arXiv: 2105.11584 [astro-ph.EP].

3.1 Introduction

Expanding and characterizing the population of known exoplanets with measured masses and orbital periods is crucial to painting a more complete picture of planet formation and evolution. A census of diverse exoplanets sheds light on worlds radically different from Earth and can provide insight into how these planets—and those orbiting the Sun—formed. Ground-based radial velocity (RV) surveys measure the Doppler shifts of stellar spectra to discover exoplanets and characterize their orbits and masses. These surveys have provided landmark discoveries that shaped our understanding of the formation and architectures of other worlds (e.g., Mayor and Queloz, 1995; Marcy, Butler, Fischer, et al., 2002; Tamuz et al., 2008).

Doppler planet searches take time to accumulate the time series measurements that trace out planetary orbits. The Keck Planet Survey (Cumming et al., 2008) used eight years of RVs from Keck-HIRES (Vogt, Allen, et al., 1994) to make the first broad measurement of giant planet occurrence ($M \sin i \geq 0.1 M_J$). This survey discovered an increase in the abundance of giant planets for orbits near the water-ice line and found that about 10% of Sun-like stars have giant planets with a semi-major axes of <3 au. The survey only reported planet detections for orbital periods shorter than 2000 days, the observational baseline of the survey. Extrapolating based on the detection of partial orbits, Cumming et al., 2008 estimated that ∼20% of such stars have a giant planet orbiting within 20 au.

Other teams of astronomers have surveyed the Northern and Southern skies in parallel with the Keck search. Mayor, Marmier, et al., 2011 used 8 years of precise HARPS RVs supplemented by additional RVs from CORALIE to measure occurrence patterns in the population of giant planets that are similar to those described above. They found that the planet mass function is “bottom heavy”. That is, low-mass planets (0.3–30 $M_\oplus$) are significantly more common than giant planets, a finding consistent with measurements from Keck Observatory by Howard, Marcy, Johnson, et al., 2010. Since then, the HARPS team has continued to discover increasingly longer-period and lower-mass planets (Udry et al., 2017; Rickman et al., 2019). Two other 'legacy' planet searches have contributed significantly to our knowledge of giant planets. Wittenmyer, Wang, et al., 2020 used data from a subset of the stars surveyed by the 18-year Anglo-Australian Planet Search, which has also uncovered a number of cold giant planets (Wittenmyer, Horner, et al., 2017; Kane et al., 2019), to measure a significant increase in giant planet occurrence at $\sim$1 au and a constant occurrence for orbits in the range $\sim$1–6 au. Similarly, the McDonald Observatory planet search has been operating for more than 20 years using the 2.7-m Harlan J. Smith Telescope, and has contributed valuable discoveries of long-period giant planets (e.g., Robertson, Endl, et al., 2012; Endl, Brugamyer, et al., 2016; Blunt et al., 2019).

We are now in the fourth decade of Doppler planet searches. As we begin to discover planets with orbital periods comparable to Saturn's, we can answer questions that require a rigorous accounting of giant planets spanning a large range of orbital distances. What is the mass versus semi-major axis distribution of planets out to 10 au? How abundant are cold gas giants beyond the water-ice line, and what can this abundance tell us about planet formation across protoplanetary disks?

The California Legacy Survey (CLS, Rosenthal et al. 2021) is uniquely suited for this work. As an unbiased radial velocity survey of 719 stars over three decades, the CLS is an excellent sample for a variety of occurrence measurements, particularly for cold gas giants. In this chapter, we explore giant planet occurrence as a function of orbital separation. In Section 2, we review the star and planet catalog of the California Legacy Survey. In Section 3, we describe our methods for computing planet occurrence. Section 4 describes the patterns of planet occurrence that we observe in the sample. In Section 5, we discuss our findings and their context. We summarize our work in Section 6.

3.2 Survey Review

The California Legacy Survey is a Doppler search for planets orbiting a well-defined sample of nearby FGKM stars conducted by the California Planet Search team (CPS; Howard, Johnson, Marcy, Fischer, Wright, Bernat, et al., 2010). The first chapter in this associated thesis describes the CLS in detail, including the stellar sample, the search methodology, and the resulting planet sample upon which this chapter and the rest of this thesis build. The CLS stellar sample was selected specifically to make the measurements reported here—planet occurrence measurements, especially of giant planets with orbits out to 10 au and beyond—and it approximates a random sample of nearby stars. In particular, stars were selected for CLS observations independent of whether planets were known to orbit them. Stars were also selected independent of their metallicity or other factors that might make them more or less likely to harbor planets.

CLS builds on Doppler measurements from the Keck Planet Search (Cumming et al., 2008), a touchstone Doppler survey of 585 stars observed with HIRES at the W. M. Keck Observatory during 1996–2004. We continued to observe those stars and an additional 134 stars at Keck Observatory through 2020. CLS also includes observations of a subset of these stars made with the Hamilton spectrometer at Lick Observatory during 1988–2011, high-cadence Keck-HIRES observations of 235 magnetically inactive stars as part of the Eta-Earth Survey (Howard, Marcy, Johnson, et al., 2010), and high-cadence Lick-APF observations of 135 of those stars (Fulton, Howard, et al., 2016; Hirsch et al., 2021). The average star has been observed for 22 years and has 71 RVs with a precision of $\sim$2 m s^{-1}. While these stars do not have homogeneous observing histories, our search methodology accounts for this by incorporating the search completeness of each star's individual dataset. (A Doppler survey that is completely homogeneous in the number, precision, and temporal spacing of measurements is infeasible given the three decade history of this planet search—indeed, this survey spans an era longer than the time during which extrasolar planets orbiting Sun-like stars have been known!) By the metric of "Doppler survey étendue" (number of stars surveyed $\times$ typical time series duration), CLS is the largest planet search to date at the $\sim$m s^{-1} level.

Our search methodology (described in Rosenthal et al. 2021) involves an automated, iterative periodogram-based search for Keplerian signals with uniform vetting to identify false positives. This methodology detected 177 planets orbiting the 719 stars in the CLS stellar sample. The algorithm is sensitive to orbital periods much longer than the baseline of our dataset, with the longest period signals detected as partial orbits.

The search was also sensitive to orbital segments only seen as linear and parabolic trends in an RV time series. There were only six such detections in our sample of trends that are not associated with known stellar binaries and are potentially consistent with planetary mass companions. Thus, nearly all orbital signals were resolved or partially resolved as Keplerian signals.

To characterize survey completeness for each star in the survey, we conducted injection-recovery tests of synthetic Doppler planet signals over a range of injected masses, orbital periods, and orbital geometries. Detected planets and CLS survey completeness are shown in Figure 3.1. We refer the reader to Rosenthal et al. (2021) for the full stellar sample and planet catalog.

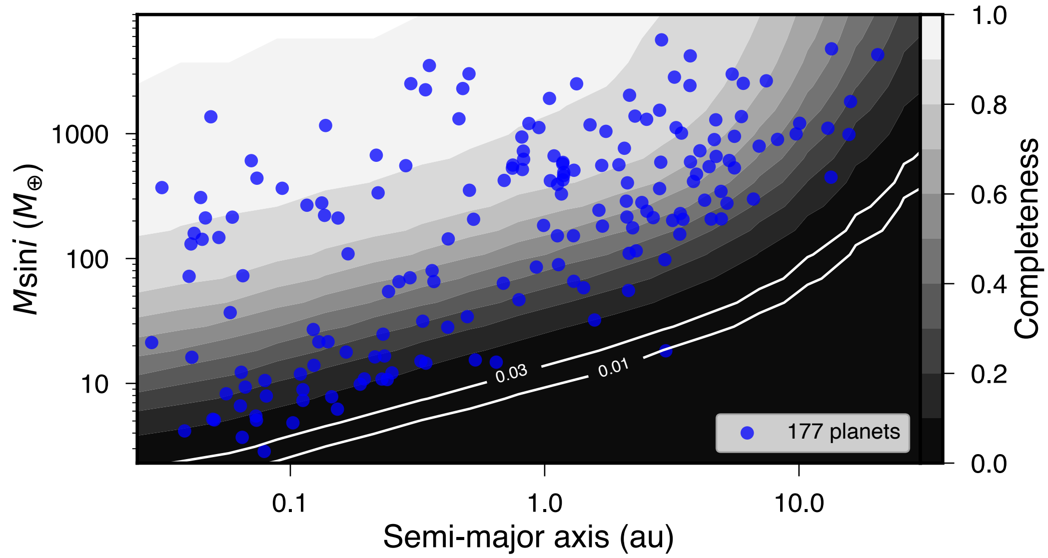

Figure 3.1: California Legacy Survey planet catalog and survey-averaged search completeness contours in semi-major axis and $M \sin i$. 3% and 1% search completeness contours are highlighted in white.

The CLS stellar sample has a median metallicity of [Fe/H]= 0.0 dex, a median stellar mass of 1.0 $M_\odot$, and a small number of evolved stars (subgiants). These are good heuristics for verifying that we successfully constructed a blind occurrence survey, since a bias toward known giant planet hosts could manifest as a metal-rich sample (Fischer, Laughlin, et al., 2005; Santos et al., 2004), a particularly massive sample, or an excess of evolved stars (Johnson, Bowler, et al., 2010).

3.3 Methods

The primary goal of this work is to measure planet occurrence. Many studies of RV or transit surveys use the intuitive occurrence measurement method known as "inverse detection efficiency" (Howard, Marcy, Bryson, et al., 2012; Petigura, Howard, et al., 2013). According to this procedure, one estimates occurrence in a region of parameter space by counting the planets found in that region, with each planet weighted by the local search completeness. One can measure the search completeness map of a survey by injecting many synthetic signals into each dataset and computing the fraction of signals in a given region that are recovered by the search algorithm in use. Inverse detection efficiency is actually a specific case of a Poisson likelihood method, in which one models an observed planet catalog as the product of an underlying Poisson process and empirical completeness map (Foreman-Mackey, Hogg, and Morton, 2014). This can be done with a parametric occurrence rate density model, like a broken power law, or a non-parametric density model, with a piecewise-constant step function. In this chapter, we used the Poisson likelihood method to model the occurrence of giant planets, taking measurement uncertainty into account.

We used the hierarchical Bayesian methodology outlined in Hogg et al., 2010 and Foreman-Mackey, Hogg, and Morton, 2014 to evaluate our occurrence likelihood. Given an observed population of planets with orbital and $M \sin i$ posteriors $\{\boldsymbol{\omega}\}$ and associated survey completeness map $Q(\boldsymbol{\omega})$, and assuming that our observed planet catalog is generated by a set of independent Poisson process draws, we evaluated a Poisson likelihood for a given occurrence model $\Gamma(\boldsymbol{\omega}|\boldsymbol{\theta})$, where Γ is an occurrence density $\frac{\mathrm{d}^2 N}{\mathrm{d}\ln(a)\mathrm{d}\ln(M \sin i)}$ and $\boldsymbol{\theta}$ is a vector of model parameters. The observed occurrence $\hat{\Gamma}(\boldsymbol{\omega}|\boldsymbol{\theta})$ of planets in our survey can be modeled as the product of the measured survey completeness and an underlying occurrence model,

$$\hat{\Gamma}(\boldsymbol{\omega}|\boldsymbol{\theta}) = Q(\boldsymbol{\omega})\Gamma(\boldsymbol{\omega}|\boldsymbol{\theta}). \tag{3.1}$$

The Poisson likelihood for an observed population of objects is

$$\mathcal{L} = e^{-\int \hat{\Gamma}(\boldsymbol{\omega}|\boldsymbol{\theta})\, d\boldsymbol{\omega}} \prod_{k=1}^{K} \hat{\Gamma}(\boldsymbol{\omega}_k|\boldsymbol{\theta}), \tag{3.2}$$

where K is the number of observed objects, and $\boldsymbol{\omega}_k$ is a vector of parameters that completely describe the kth planet's orbit. In our case, the two relevant parameters are $M \sin i$ and semi-major axis a, taken from the broader set that includes eccentricity, time of inferior conjunction, and argument of periastron. The Poisson likelihood can be understood as the product of the probability of detecting an observed set of objects (the product term in Equation 2) and the probability of observing no additional objects in the considered parameter space (the exponentiated integral). Equations 1 and 2 serve as the foundation for our occurrence model but do not take into account uncertainty in measurements of planetary orbits and minimum masses. In order to do this, we used `RadVel` and `emcee` to empirically sample the orbital posteriors of each system (Fulton, Petigura, Blunt, et al., 2018; Foreman-Mackey, Hogg, Lang, et al., 2013a). We hierarchically modeled the orbital posteriors of each planet in our catalog by summing our occurrence model over many posterior samples for each planet. The hierarchical Poisson likelihood is therefore approximated as

$$\mathcal{L} \approx e^{-\int \hat{\Gamma}(\boldsymbol{\omega}|\boldsymbol{\theta})\, d\boldsymbol{\omega}} \prod_{k=1}^{K} \frac{1}{N_k} \sum_{n=1}^{N_k} \frac{\hat{\Gamma}(\boldsymbol{\omega}_k^n|\boldsymbol{\theta})}{p(\boldsymbol{\omega}_k^n|\boldsymbol{\alpha})}, \tag{3.3}$$

where N_k is the number of posterior samples for the kth planet in our survey and $\boldsymbol{\omega}_k^n$ is the nth sample of the kth planet's posterior. $p(\boldsymbol{\omega}|\boldsymbol{\alpha})$ is our prior on the individual planet posteriors. We placed linear-uniform priors on $M\sin i$ and log-uniform priors on a. We used `emcee` to sample our hierarchical Poisson likelihood.

We used two different occurrence frameworks to model our planet population. The first is a non-parametric model across bins uniformly spaced in $\ln(M\sin i)$ and $\ln(a)$, with a set of steps $\boldsymbol{\Delta}$ of height $\boldsymbol{\theta}$. We define this framework with the occurrence function

$$\Gamma_N(\boldsymbol{\omega}|\boldsymbol{\theta}) = \theta_n | \boldsymbol{\omega} \in \Delta_n. \tag{3.4}$$

The second framework is a broken power law as a function of semi-major axis, defined with the function

$$\Gamma_B(a|C, \beta, a_0, \gamma) = C(a/\mathrm{au})^{\beta}(1 - e^{-(a/a_0)^{\gamma}}), \tag{3.5}$$

where C is a normalization constant, β is the occurrence power law index beyond the breaking point, a_0 determines the semi-major axis location of the breaking point, and $\beta + \gamma$ is the power law index within the breaking point. This model assumes a giant planet mass function that does not change with respect to semi-major axis. We fit this model to our population in order to explore whether giant planet occurrence falls off beyond the water-ice line.

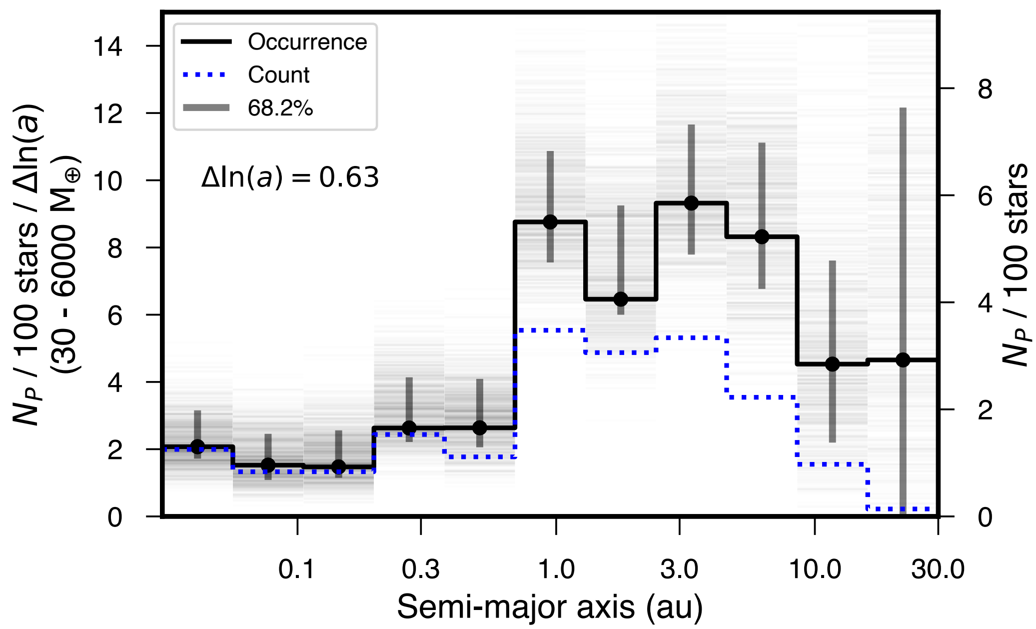

Figure 3.2: Non-parametric occurrence rates for semi-major axes of 0.03–30 au for planets with minimum masses from 30–6000 $M \sin i$ assuming uniform occurrence across $\ln(M \sin i)$. The dashed blue line represents a planet count in each semi-major axis bin without correcting for completeness; bold lines and dots show the maximum posterior values for the Poisson likelihood model; vertical lines represent 15.9–84.1% confidence intervals (except for the last bin, which is not separated from zero and shows 0–68.2%); and transparent steps show draws from the occurrence posterior. We see a clear enhancement around 1–10 au, and a tentative falloff beyond that range.

3.4 Results

Enhancement for giant planets

Figure 3.2 shows occurrence rates as a function of semi-major axis for planets with masses between 30 $M_\oplus$ and 6000 $M_\oplus$, derived using the non-parametric model described in §3.3 and assuming uniform occurrence across $\ln(M \sin i)$. We confirmed the previous result from Wright, Upadhyay, et al., 2009, Cumming et al., 2008, Fernandes et al., 2019, and Wittenmyer, Wang, et al., 2020 that giant planet occurrence is enhanced by a factor of four beyond 1 au compared to within 1 au. Specifically, planets more massive than 30 $M_\oplus$ are 2–4 times more common at orbital distances between 1–3 au relative to 0.1–0.3 au. Using our broken power law model, we find a median power law slope inside the break of $0.72^{+0.16}_{-0.20}$, which is 2 σ higher than the power law slope measured by Cumming et al., 2008 (0.26±0.1).

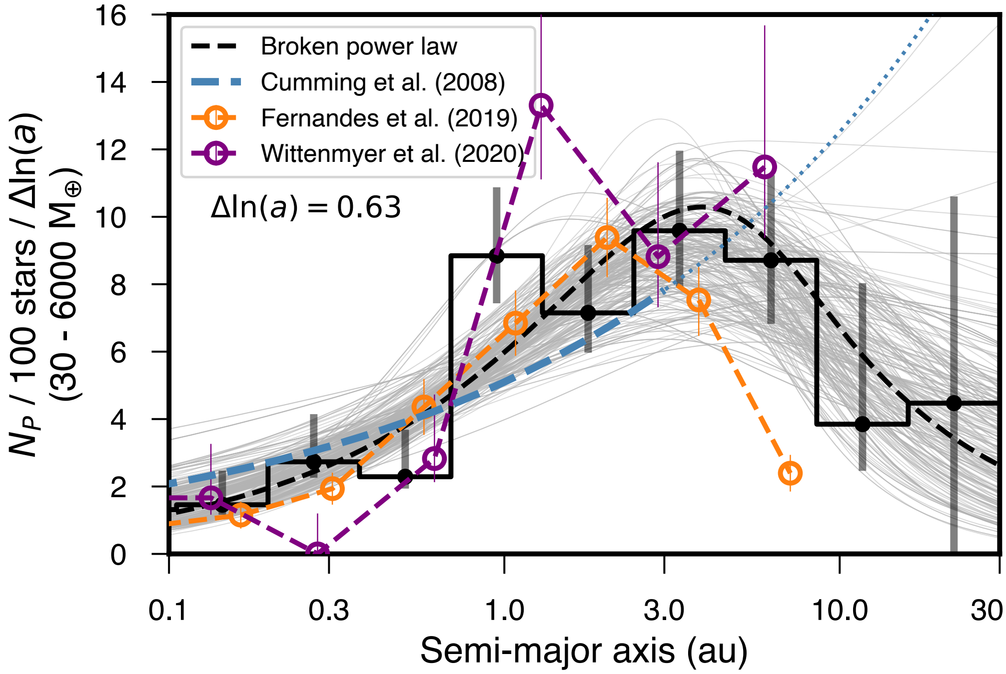

Figure 3.3: Our broken power law model, juxtaposed with our non-parametric model and measurements from Fernandes et al., 2019 and Wittenmyer, Wang, et al., 2020. The transparent curves represent draws from the broken power law posterior. We find that the power law index beyond the break is ~2.5σ-separated from zero, implying an occurrence falloff beyond the water-ice line. Cumming et al. (2008) performed a power-law fit to the occurrence rates of planets orbiting only within 3 au; the light dotted blue line represents an extrapolation to wider separations.

This difference is likely caused by the single power law model being pulled to lower values due to neglecting a flattening or turnover in occurrence at long orbital periods since Cumming et al., 2008 was limited to planets orbiting inside 3 au.

Distribution of giant planets beyond 3 au

Due to low completeness beyond our observational baselines, our occurrence results beyond 10 au are highly uncertain. However, we can estimate occurrence trends with the broken power law model described in §3.3. Figure 3.3 shows the broken power law results juxtaposed with the non-parametric results, and Figure 3.4 presents the posteriors for the parametric model parameters. The medians and 68th percentile credible intervals for the broken power law model are listed in Table 3.1. Both assume uniform occurrence across $\ln(M \sin i)$. We find that 99.4% of the posterior samples are consistent with a plateauing or declining occurrence rate beyond a peak around $3.6^{+2.0}_{-1.8}$ au. We find that the power law index beyond the peak is

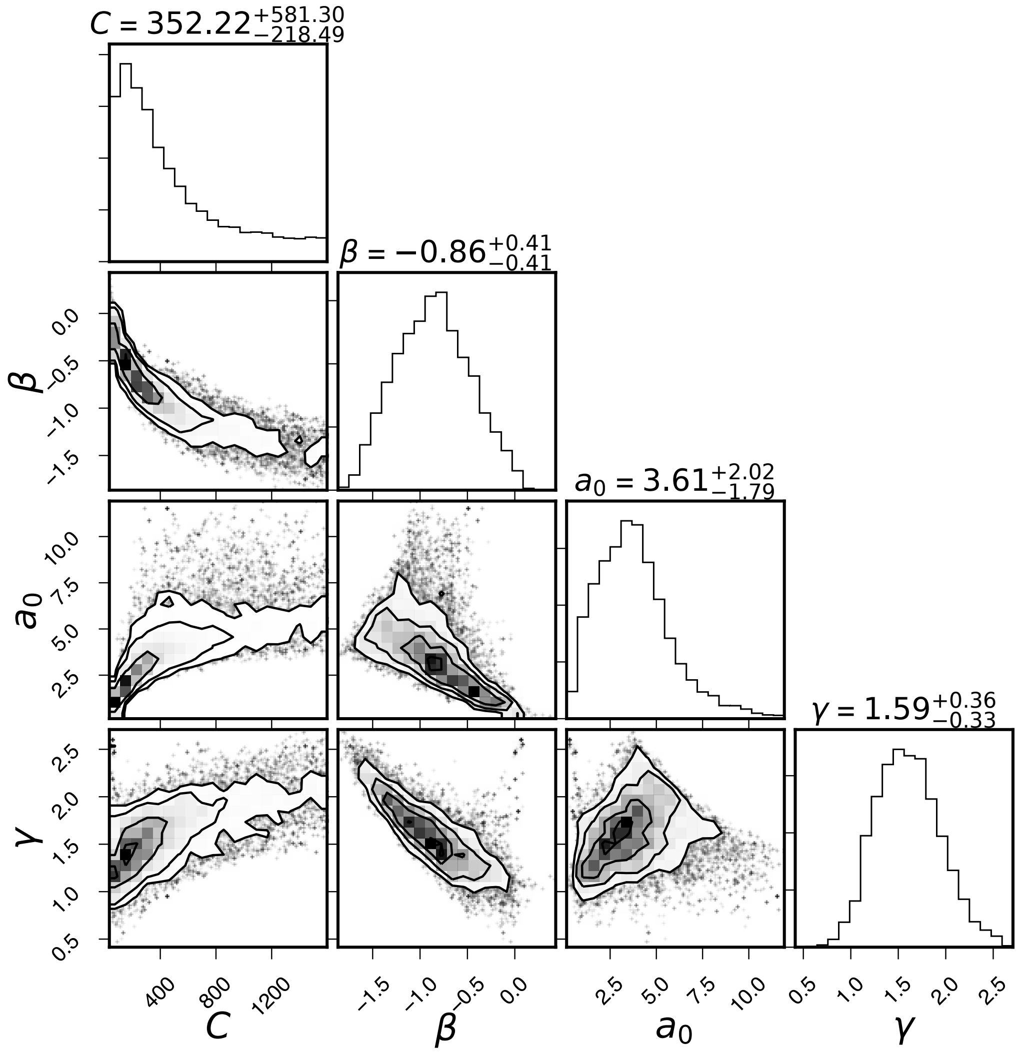

Figure 3.4: Broken power law posterior. C is a normalization constant, β is the power law index beyond the break, a_0 determines the location of the break in units of au, and $\beta+\gamma$ is the power law index within the break. The index beyond the break β is ~ 99.1%-separated from zero.

$\beta = -0.86^{+0.41}_{-0.41}$. This suggests a much shallower decline relative to the estimates of Fernandes et al., 2019 but is also potentially discrepant with the constant prediction of Wittenmyer, Wang, et al., 2020, as our model still measures a falloff. The results of our non-parametric fit are less clear, with integrated occurrence rates of $14.1^{+2.0}_{-1.8}$ and $8.9^{+3.0}_{-2.4}$ giant planets per 100 stars between 2–8 au and 8–32 au respectively. This suggests a fall-off in occurrence beyond 8 au with 1.5σ confidence.

Table 3.1: Broken Power-Law Model Parameters

Parameter	Value
C	350^{+580}_{-220}
β	$-0.86^{+0.41}_{-0.41}$
a_0	$3.6^{+2.0}_{-1.8}$ au
γ	$1.59^{+0.36}_{-0.33}$

Comparing sub- and super-Jovians

Figure 3.5 compares non-parametric occurrence rates for giant planets more and less massive than 300 $M_\oplus$. We find a quantitatively similar occurrence enhancement around 1–10 au for both the sub-Jovian-mass and Jovian-mass planets. However, we lack the sensitivity to measure the occurrence rate of sub-Jovian mass planets beyond 10 au, to assess whether they exhibit the fall-off in occurrence at large orbital separations seen when examining occurrence across both mass ranges. The sub-Jovian planets are more common than the super-Jovian planets across a wide range of separations, particularly beyond the water ice line. We find a similar enhancement for sub-Saturns below 150 $M_\oplus$, implying that this occurrence enhancement is independent of planet mass.

We more concretely measured occurrence as a function of mass by performing a non-parametric fit to our sample within 1–5 au. Figure 3.6 shows occurrence as a function of $M \sin i$ within 30–3000 $M_\oplus$, in four steps. This figure shows that our assumption of a uniform $\ln(M \sin i)$ distribution beyond the ice line is valid up to 900 $M_\oplus$, but the distribution falls off with $\sim 2\sigma$ significance above 900 $M_\oplus$. If this is also true beyond 5 au, where low completeness prevents us from making a similar measurement, then we may be underestimating broad giant planet occurrence in our lowest-completeness region of parameter space, beyond 10 au. This is because our only detections in that regime are more massive than 300 $M_\oplus$, and all but one of them are more massive than 900 $M_\oplus$.

Occurrence with respect to stellar mass and metallicity

In addition to measuring occurrence with respect to semi-major axis and $M \sin i$ we measured the broad occurrence rate of giant planets more massive than 100 $M_\oplus$ and within 1–5 au with respect to host-star mass and metallicity. We chose a lower limit of 100 $M_\oplus$ instead of 30 $M_\oplus$ in order to restrict our analysis to search-complete regions within 1–5 au, since 30 $M_\oplus$ planets are effectively undetectable beyond

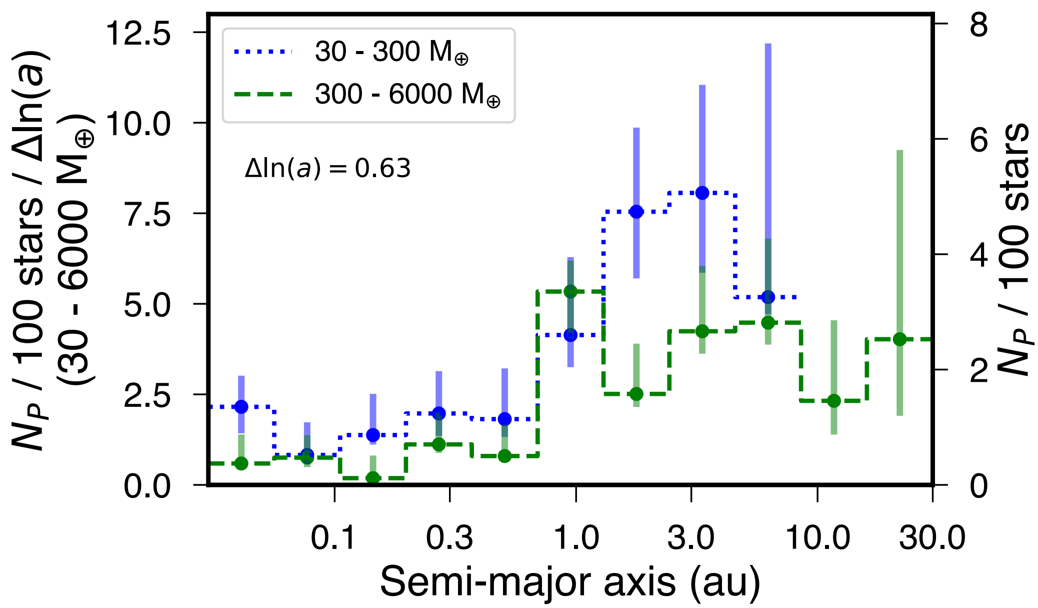

Figure 3.5: A comparison between sub- and super-Jovian occurrence. Steps and dots show maximum posterior values, and vertical lines show 15.9–84.1% confidence intervals. The sub-Jovians are consistently more common than the super-Jovians, and both populations are enhanced beyond 1 au. Combining these two populations produces the same trends seen when we assume uniform occurrence across all masses.

3 au. For each of these two stellar properties, we computed occurrence across six divisions, in steps of 0.2 $M_\odot$across 0.3–1.5 $M_\odot$and 0.15 dex across -0.5–0.4 dex respectively. Figure 3.7 shows occurrence with respect to host-star mass and with respect to host-star [Fe/H]. Both of our measurements agree with prior results. Johnson, Aller, et al., 2010, whose stellar sample was excluded from CLS due to its bias toward giant planet hosts, measured giant planet occurrence across stellar mass and found an increase in occurrence with increasing stellar mass beginning near 1 $M_\odot$. Wittenmyer, Butler, et al., 2020 independently found an increase in giant planet occurrence beyond 1 $M_\odot$. We see the same phenomenon in our sample, as presented in the left panel of Figure 3.7. Similarly, Fischer, Laughlin, et al., 2005 found that giant planet occurrence increases with increasing [Fe/H] beyond 0.1 dex, as did Reffert et al., 2015 and Jones et al., 2016. We see the same transition near 0.1 dex in the right panel of Figure 3.7.

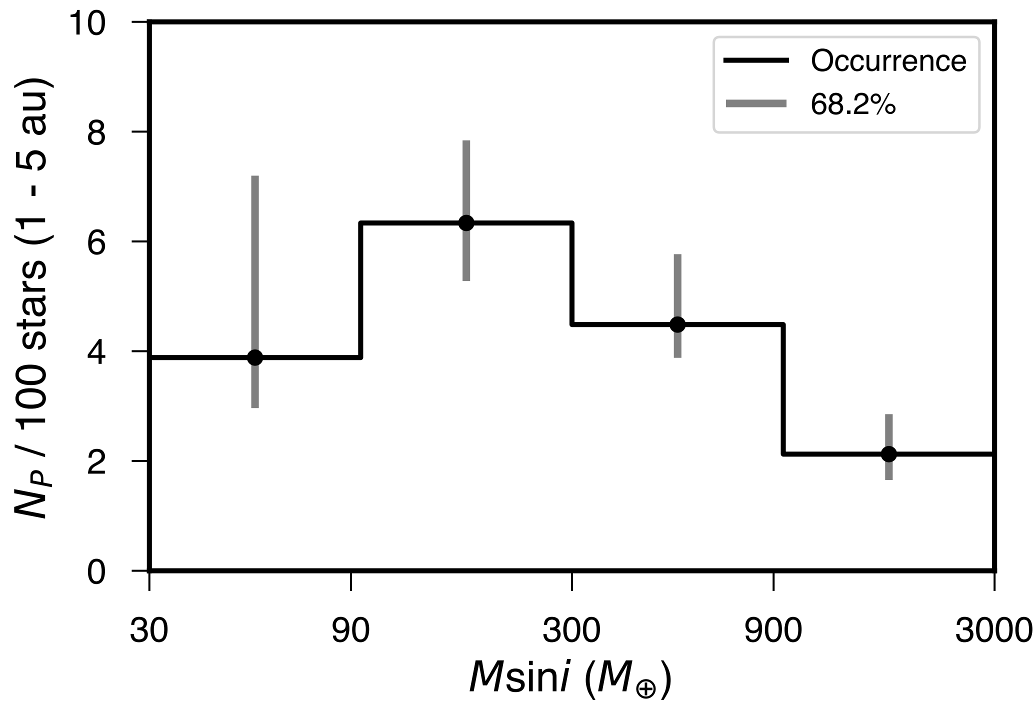

Figure 3.6: Planet occurrence within 1–5 au with respect to $M \sin i$. Steps and dots show maximum posterior values, and vertical lines show 15.9–84.1% confidence intervals. The mass function is constant within 30–900 $M_\oplus$, and falls off beyond 900 $M_\oplus$.

3.5 Discussion

Comparison to previous RV surveys

The last few years have seen a number of RV studies examining the population of long-period planets. Fernandes et al., 2019 probed planet occurrence as a function of orbital period by extracting planetary minimum masses and periods, as well as completeness contours, from a catalog plot shown in Mayor, Marmier, et al., 2011, which presented a HARPS (Mayor, Pepe, et al., 2003) and CORALIE (Baranne et al., 1996) blind radial velocity survey of 822 stars and 155 planets over 10 years (corresponding to a 4.6 au circular orbit around a Solar-mass star). Mayor, Marmier, et al., 2011, who did not publish their HARPS and CORALIE RVs, measured giant planet occurrence as a function of orbital period out to 4000 days, in the range of the water-ice line. Fernandes et al., 2019 pushed out to low-completeness regimes and estimated a sharp falloff in occurrence beyond the water-ice line. They measured an integrated occurrence rate of 1.44±0.54 giant planets (0.1–20 M_J) per 100 stars for separations between 3.8 and 7.1 au. Our results indicate a much higher occurrence

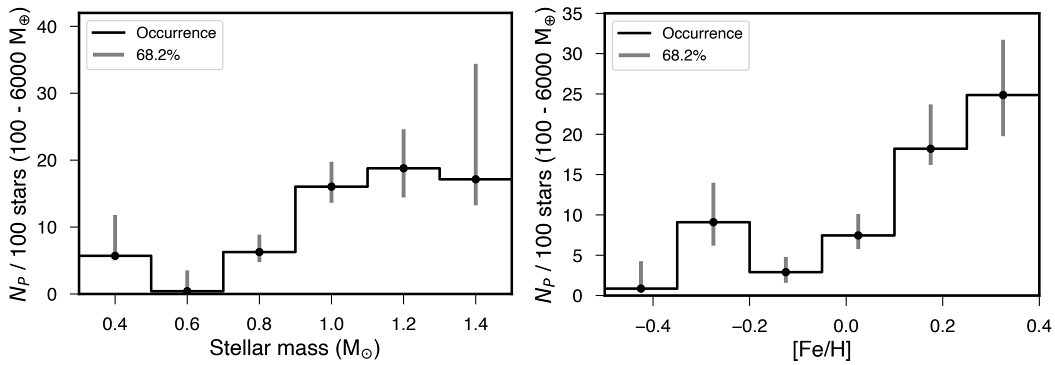

Figure 3.7: Left: occurrence of giant planets more massive than 100 $M_\oplus$ and within 1–5 au as a function of host star mass, in six splits. Steps and dots show maximum posterior values, and vertical lines show 15.9–84.1% confidence intervals. There is an increase in occurrence beyond roughly 1 $M_\odot$, which is in agreement with Johnson, Aller, et al., 2010's original measurement of giant planet occurrence versus host-star mass. Right: occurrence of giant planets more massive than 100 $M_\oplus$ and within 1–5 au as a function of host star metallicity, in six splits. Steps and dots show maximum posterior values, and vertical lines show 15.9–84.1% confidence intervals. There is a clear increase in occurrence beyond roughly 0.1 dex, which is in agreement with Fischer, Laughlin, et al., 2005's original report of a correlation between giant planet occurrence and host-star metallicity.

rate for the same planets at those separations; $15.5^{+3.2}_{-3.0}$ giant planets per 100 stars. The treatment of partial orbits in Mayor, Marmier, et al., 2011 is unclear, and they only measured occurrence with respect to orbital period out to 3000 days (∼4 au). If Mayor, Marmier, et al., 2011 under-reported partial orbits beyond this period in their sample or overestimated sensitivity to partial orbits, then that could explain the large discrepancy between this work and Fernandes et al. (2019) at separations beyond 10 au.

In contrast, Wittenmyer, Wang, et al., 2020, which drew from the Anglo-Australian Planet Search (Tinney et al., 2001) to construct a blind survey of 203 stars and 38 giant planets over 18 years, found that giant planet occurrence is roughly constant beyond the water-ice line, out to almost 10 au. Wittenmyer, Wang, et al., 2020 reports an occurrence rate of $6.9^{+4.2}_{-2.1}$ giant planets > 0.3 M_J per 100 stars with periods between 3000 and 10,000 days (≈4–9 au). Our integrated occurrence rate in the same region of parameter space is slightly higher at $12.6^{+2.6}_{-2.0}$ giant planets per 100 stars but it is consistent to within 1 σ with the Wittenmyer, Wang, et al., 2020 result.

Comparison to Kepler survey

Foreman-Mackey, Morton, et al., 2016 performed an automated search for long-period transiting exoplanets in a set of archival *Kepler* light curves of G and K stars. For planets between 1.5–9 au and 0.01–20 M_J and using a probabilistic mass-radius relationship drawn from Chen and Kipping, 2016, they found an occurrence rate density of $\frac{d^2N}{d\ln(a)d\ln(M)} = 0.068 \pm 0.019$. We applied our occurrence model to the same parameter space and found $\frac{d^2N}{d\ln(a)d\ln(M \sin i)} = 0.0173 \pm 0.0022$. The *Kepler* measurement is 2.65σ separated from our measurement. We are far less sensitive to planets in the 0.01–0.1 M_J regime than Foreman-Mackey, Morton, et al., 2016; this may partly explain the discrepancy in our results.

Comparison to direct imaging surveys

RV surveys have recently begun to approach baselines long enough to detect and place limits on the frequency of planets like those detected by direct imaging. One caveat is that direct imaging surveys usually target stars younger than 100 Myr, while RV surveys generally target stars older than 1 Gyr. Young planets retain significant heat from their formation and are bright in the infrared wavelengths covered by direct imaging surveys. However, young stars also tend to be active and rapidly rotating, which makes precise RV work difficult. Because of this, there is minimal overlap between planets that have been detected by direct imaging and planets that have been detected by radial velocity measurements.

We can still compare rates across these detection methods by making the assumption that giant planet occurrence does not change as host stars age beyond ~10 Myr, once protoplanetary disks have dissipated. We compared our occurrence model to the results of two direct imaging surveys of nearby stars. Biller et al., 2013 imaged 80 stars in nearby moving groups and detected a small number of brown dwarf companions but no planetary-mass companions. They used stellar evolution and planet formation models to estimate constraints on cold giant occurrence from their nondetections and sensitivity. More recently, Nielsen et al., 2019 imaged 300 stars and detected six planets and three brown dwarfs. Figure 3.8 compares these results to our occurrence measurements in their respective regions of parameter space. Our measurements are compatible with the limits placed on planets with masses 1–20 M_J and separations between 10–50 au by Biller et al., 2013, depending on their assumed stellar evolutionary model that determines the expected brightness of young giant planets. Our measurement for planets with masses 5–14 M_J orbiting between 10–100 au is in excellent agreement with the results of Nielsen et al., 2019. The only shared quality of our modeling methods is a Poisson counting likelihood. With the caveat of small number statistics, this is a remarkable benchmark for comparing exoplanet occurrence across independent search methods.

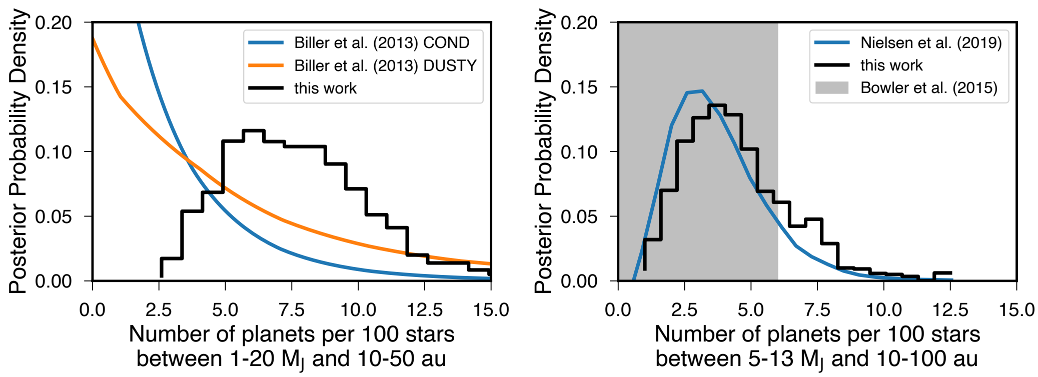

Figure 3.8: Occurrence rate comparison to direct imaging studies. *Left*: Frequency of cool, massive companions with the direct imaging study of Biller et al. (2013). While they did not detect any planets in their survey they were able to put upper limits on the frequency of companions using assumptions of either hot-start (COND) or cold-start (DUSTY) models for planetary formation and infrared brightness. *Right*: Same as left, but compared with the results of Bowler, Liu, et al. (2015) and Nielsen et al. (2019) for the mass and separation limits specified in the x-axis label. The gray shading represents the 95% upper limit on occurrence from Bowler, Liu, et al. (2015).

Comparison to gravitational microlensing surveys

We compare our model to the microlensing surveys of Cassan et al., 2012 and Clanton and Gaudi, 2016. Like all gravitational lensing surveys, these studies assume a broad prior for stellar type based on Galactic observations, a prior that peaks in the M dwarf range. Our planet-hosting stars have a much higher median mass than this range, but since the gravitational lensing estimates comes purely from a galactic model prior, we chose to perform this broad comparison across stellar masses with the knowledge that the mass range for the lensing numbers is poorly constrained. Figure 3.9 shows that our estimates agree with broad constraints from the pure lensing survey (Cassan et al., 2012). On the other hand, the constraints of Clanton and Gaudi, 2016 strongly disagree with our planet occurrence measurement in the same parameter box. This may be due to that study having a significantly better constrained sample of M dwarfs, which would separate their stellar sample from our broader FGKM sample. Endl, Cochran, et al., 2006, Bonfils et al., 2013, and Montet et al., 2014 performed independent RV surveys of M dwarfs and all

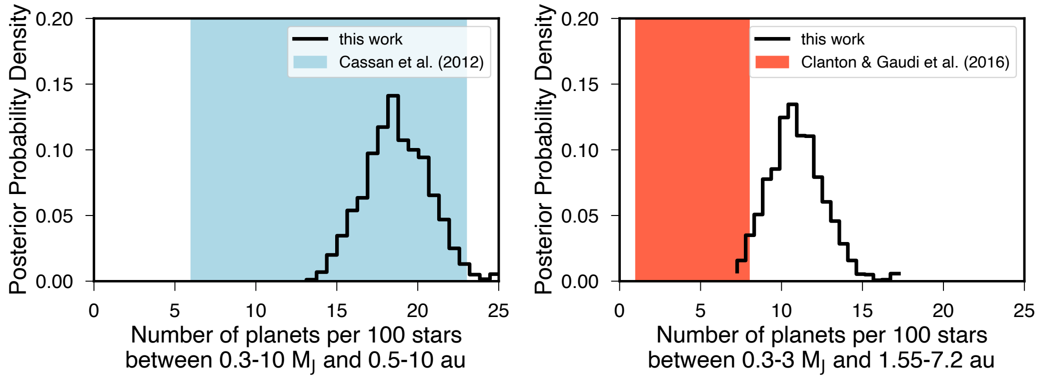

Figure 3.9: *Left:* Occurrence rate comparison with the microlensing survey of Cassan et al. (2012). We plot the 1 σ limits from Cassan et al. (2012) as the shaded blue region. The occurrence rate posterior from this work is plotted in black. *Right:* Occurrence rate comparison with the combined analysis of Clanton and Gaudi (2016). The occurrence rate posterior from this work is plotted in black. The 1 σ limits from Clanton and Gaudi (2016) are indicated by the shaded red region. Clanton and Gaudi (2016) combine constraints from direct imaging, microlensing, and previous radial velocity studies.

showed that M dwarfs have a significantly lower giant planet occurrence rate than more massive stars. This implies that a survey of M dwarfs should yield a lower giant planet occurrence rate than a broad survey of FGKM stars, and this is exactly what we see in our comparison to Cassan et al., 2012.

Implications for planet formation

Cumming et al. (2008) first identified an enhancement in the occurrence rate of giant planets beyond orbital periods of ∼300 days. We expect such enhancements based on planetary migration models (Ida and Lin, 2004). The orbital period distribution in Cumming et al., 2008 predicted a smooth rise in occurrence toward longer orbital periods, but we observed a sharp transition around 1 au, as seen in Figure 3.2. Ida and Lin (2008b) later suggested that additional solid materials due to ices in the protoplanetary disk could augment the formation of gas giant planets and cause a rapid rise in the occurrence rate of these planets beyond the water-ice line.

If increased solids near and beyond the ice line cause a sharp rise in the occurrence rate, then we might expect this rise to be more well-defined when looking in a unit more closely related to the temperature in the protoplanetary disk. In Figure 3.10, we plot the occurrence rate as a function of stellar light intensity relative to Earth. The occurrence rate with respect to flux is qualitatively similar to the rate with respect to orbital separation. We do not see strong evidence that the occurrence rate enhancement is any more localized in terms of stellar light intensity relative to Earth.

We can separate the puzzle of gas giant formation into two components: the growth of solid cores that are large enough to undergo runaway gas accretion, and the process of gas accretion onto solid cores. It is currently unclear whether giant planet occurrence increases beyond the ice line because cores form more easily in this region, or because conditions are more favorable for rapid gas accretion onto solid cores. A number of studies (e.g. Morbidelli et al., 2015; Schoonenberg and Ormel, 2017; Drążkowska and Alibert, 2017) have argued that that core growth should be enhanced beyond the ice line. If the solid grain sizes and densities beyond the ice line are enhanced during the earliest stages of planet formation, it would facilitate pebble clumping that leads to planetesimal formation and also result in higher pebble accretion rates onto the growing core (e.g. Bitsch, Raymond, et al., 2019).

It is also possible that gas giants are more common beyond the ice line because it is easier for cores to rapidly grow their gas envelopes in this region. The rate at which the growing planet's envelope can cool and contract (hence accreting more gas) depends sensitively on its envelope opacity (e.g. Bitsch and Savvidou, 2021). In a recent study, Chachan, Lee, et al., 2021 used dust evolution models to study the effect of dust opacity and dust-to-gas ratio on giant planet formation in the epoch immediately following the end of core formation. They found that as the disk evolves, decreasing dust opacity beyond the water-ice line allows for higher gas accretion rates in this region.

Ida, Tanaka, et al. (2018) recently updated their models with an improved treatment of Type II migration. This mechanism would produce a broad semi-major axis distribution with many giant planets migrating inward to separations less than 1 au. However, Fernandes et al. (2019) show that this model does not agree well with the occurrence contrast between the peak and the inner regions of these systems. Our results are in close agreement with those of Fernandes et al. (2019) for separations less than 3 au. The multi-core accretion models of Mordasini (2018) are also in good agreement with the overall shape of the semi-major axis distribution, but they underestimate the absolute occurrence rate of giant planets. This could be due to the finite number of cores injected into their simulations.

One common theme among planet formation models of gas giants is that protoplanets

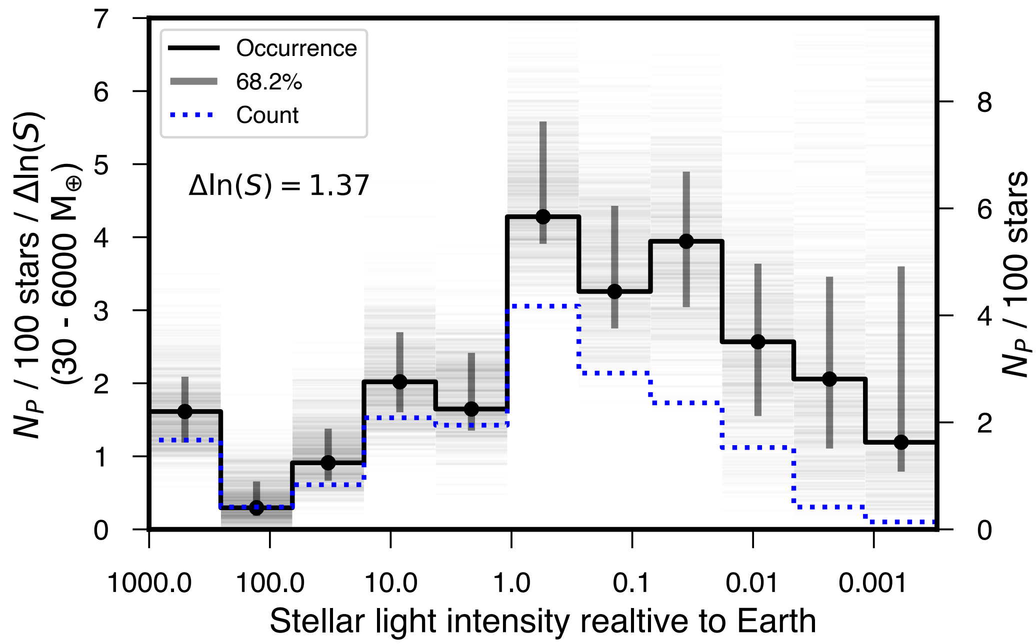

Figure 3.10: Analogous to Figure 2, occurrence with respect to stellar light intensity instead of orbital separation. Here we see a similar enhancement in the occurrence rate of giant planets where the insolation flux is equal to that of Earth and tentative evidence for a fall off in occurrence just beyond that.

tend to migrate inward, all the way to the inner edge of the disk, on timescales much shorter than the gas dissipation timescale. This tends to produce an enhancement of occurrence closer to the star and/or many planets being engulfed by the host star. Jennings et al. (2018) attempt to solve this issue by simultaneously modeling the effects of photoevaporation and viscous evolution on the gas disk. They find that, depending on the dominant energy of the evaporating photons, this could clear gaps in the disk that halt Type I migration and creates a pile-up of planets at orbital separations between 0.8–4 au. They showed that this can produce very strong and narrow enhancements near certain orbital separations, but it is conceivable that the shape of the final semi-major axis distribution would actually be driven by the spectral energy distributions of host stars during the early years of their formation.

Hallatt and Lee (2020) also proposed gap formation in the protoplanetary disk shortly after the formation of gas giant planets as a mechanism to slow or halt migration at preferred orbital separations. Their model requires that the giant planets that form further out in the disk are more massive in order to reproduce the observed enhancements. We expect this to be the case if the dust content of disk envelopes is very low.

The observed enhancement in the occurrence rate of sub-Jovian planets near 1–10 au, seen in Figure 3.5, suggests that the processes that drive the formation and pile-up of planets at those orbital distances also apply to these lower-mass planets. It appears just as likely for a gaseous planet to undergo runaway accretion and grow into a Jovian planet as it is to halt that runaway accretion process early and remain in the sub-Saturn regime.

Unfortunately, it is difficult to extract significant constraints on planet formation models from semi-major axis distributions alone. Future planet catalogs produced by Gaia and The Roman Space Telescope will help to measure the precise shape of the occurrence enhancement around 1 au with planet samples several orders of magnitude larger, but the stellar samples will be different from ours. We plan for future works in this series to analyze the host star metallicity, eccentricity, and multiplicity distributions of our sample, in the hopes of uncovering evidence that discriminates between different planet formation models.

3.6 Conclusion

In this work, we utilize the catalogue of stars, RV-detected planets, and completeness contours from Rosenthal, Fulton, et al., 2021 to measure giant planet occurrence as a function of semi-major axis. We applied a hierarchical Bayesian technique to incorporate measured search completeness and uncertainties in our observations into uncertainties in our occurrence rates. Our results are consistent with previous studies that have found a strong enhancement in the occurrence rates of these planets around 1 au.

We find that the occurrence of planets less massive than Jupiter ($30 \leq M \sin i \leq 300$ $M_{\oplus}$) is enhanced near 1–10 au in concordance with their more massive counterparts. We find that a fall-off in giant planet occurrence at larger orbital distances is favored over models with flat or increasing occurrence, with 2.5 σ confidence from our broken power-law model and with 1.5 σ confidence from our non-parametric model. Additionally, our occurrence measurements beyond 10 au are consistent with with those derived from direct imaging surveys.

All code used in this chapter is available at `github.com/California-Planet-Search/rvsearch` and `github.com/leerosenthalj/CLSII`. This research makes use of GNU Parallel (Tange, 2011). We made use of the following publicly available Python modules: `astropy` (Astropy Collaboration et al., 2013), `matplotlib`

(Hunter, 2007), `numpy/scipy` (Walt et al., 2011), `pandas` (McKinney, 2010), `emcee` (Foreman-Mackey, Hogg, Lang, et al., 2013a), and `RadVel` (Fulton, Petigura, Blunt, et al., 2018).

Chapter 4

ON THE SHOULDERS OF (SOME) GIANTS: THE RELATIONSHIP BETWEEN INNER SMALL PLANETS AND OUTER MASSIVE PLANETS

Rosenthal, L. J. et al. (Dec. 2021). "The California Legacy Survey III. On The Shoulders of (Some) Giants: The Relationship between Inner Small Planets and Outer Massive Planets". In: *arXiv e-prints*, arXiv:2112.03399, arXiv:2112.03399. arXiv: 2112.03399 [astro-ph.EP].

4.1 Introduction

The relationship between small, close-in planets and outer giant companions reveals much about planet formation. Gas giant interactions with protoplanetary disks can create low-density gaps that halt the inward drift of gas and solids, possibly suppressing the formation of close-in small planets (Lin and Papaloizou, 1986; Moriarty and Fischer, 2015; Ormel et al., 2017). It is also possible that warm or eccentric giants disrupt the growth of of small planets by pebble or planetesimal accretion, or destabilize the orbits of nascent small planet cores, as predicted by synthetic population studies (Bitsch, Trifonov, et al., 2020; Schlecker et al., 2020). These phenomena would lead to a population of small planets without outer giant companions, or an absence of companions within a certain range of mass, semi-major axis, and eccentricity.

On the other hand, the same stellar properties that facilitate giant planet formation, such as high metallicity (Fischer, Laughlin, et al., 2005), may also enhance small planet formation. If higher metallicity stars were more likely to form super-Earths, we would expect to see a metallicity dependence in their observed occurrence rate regardless of the presence or absence of an outer companion. Petigura, Marcy, Winn, et al., 2018 analyzed the metallicity distribution of *Kepler* stars and planet hosts and found that warm sub-Neptune (1.7–4 $R_\oplus$) occurrence is weakly correlated with host-star metallicity, doubling from -0.4 dex to +0.4 dex with $\sim 2\sigma$ significance. However, Moe and Kratter, 2019 and Kutra and Wu, 2020 later found that this correlation disappears if one decorrelates against the metallicity dependence of close binaries, which do not host short-period planets. This leaves open the possibility that the giant planets formed in metal-rich disks might directly facilitate small planet formation via their dynamical impact on the protoplanetary disk structure (e.g., Hasegawa and Pudritz, 2011; Buchhave et al., 2014).

We can explore the tension between these ideas by using exoplanet surveys to measure the conditional probability that gas giant hosts also host at least one inner small planet, and comparing that value to the overall occurrence rates of close-in small planets and distant giant planets, beyond roughly 0.3 au. If inner small planet companions to cold giants are rarities compared to the broader sample of small planets, then we can deduce that giant planets in a certain mass and semi-major axis range suppress small planet formation. Conversely, if small planet companions to cold giants are common, this implies that disks that form cold gas giants also provide favorable conditions for small planet formation, or that cold giants actively facilitate small planet formation.

Recently, Zhu and Wu, 2018 and Bryan, Knutson, Lee, et al., 2019 independently used samples of stars with known super-Earths, most of which have masses less than 10 $M_\oplus$, to directly estimate the fraction of super-Earth hosts that have outer gas giant companions. Furthermore, each analysis used Jupiter-analog occurrence rates from Jones et al., 2016 and Rowan et al., 2016 to infer the fraction of cold giants that host inner super-Earths. Bryan, Knutson, Lee, et al., 2019 found that $102^{+34}_{-51}\%$ of stars with Jupiter analogs (3–7 au, 0.3–13 M_J) also host an inner super-Earth, while Zhu and Wu, 2018 reported $90 \pm 20\%$ for the same measurement. Both studies predicted that nearly all Jupiter analogs host inner small planets, with high uncertainties in conditional probability due to the indirect nature of this Bayesian inference. So far, no study has directly measured the rate at which gas giants are accompanied by inner small planets. To make this measurement, we need a large sample of cold Jupiters with RV data sets that are also sensitive to the presence of small inner planets, which requires long-baseline radial velocity (RV) observations. The measurement must also be sensitive to small inner planets, which requires high-cadence radial velocities or coverage by a photometric transit survey. The former is a costly undertaking that can only be done over many nights of ground-based RV observing, and the latter can only detect planets with edge-on or nearly edge-on orbits.

The California Legacy Survey (CLS; Rosenthal, Fulton, et al., 2021) is uniquely well-suited for this measurement. As a blind RV survey of 719 stars over three decades, it produced a sample that is appropriate for a variety of occurrence measurements, is rich in cold giants, and contains enough stars with high-cadence observations that we have some sensitivity in the small-planet regime. In this chapter, we leverage this survey to explore the relationship between close-in small planets, which we limit to 0.03–1 au and 2–30 $M_\oplus$, and outer giant companions, which we constrain in two different ways defined below. In Section 2, we review the star and planet catalog of the California Legacy Survey. In Section 3, we describe our methods for computing planet occurrence. In Section 4, we present our results. In Section 5, we discuss our findings and their context.

4.2 Survey Review

The California Legacy Survey is a sample of 719 RV-observed FGKM stars and their associated planets created to provide a stellar and planetary catalog for occurrence studies (Rosenthal, Fulton, et al., 2021). We approximated a quantifiably complete survey by selecting 719 stars that were observed by the California Planet Search (CPS) (Howard, Johnson, Marcy, Fischer, Wright, Bernat, et al., 2010) and originally chosen without bias towards a higher or lower than average likelihood of hosting planets. We took our first observations in 1988 with the Lick-Hamilton spectrograph (Fischer, Marcy, et al., 2014), and our latest observations in 2020 with the Keck-HIRES and the Lick-APF spectrographs. Our typical observational baseline is 22 yr, and our typical RV precision is 2 m s^{-1}. We used an automated and repeatable iterative periodogram method to search for planet candidates, implemented in the open-source package `RVSearch` (Rosenthal, Fulton, et al., 2021), and performed uniform vetting to identify false positives. This left us with 178 planets in our sample, 43 planets with $M \sin i < 30\ M_\oplus$, and 135 planets with $M \sin i \geq 30\ M_\oplus$. Figure 4.1 shows our sample of small close-in planets and outer giant companions.

Our stellar sample has a median metallicity of 0.0 [Fe/H], a median stellar mass equal to 1.0 $M_\odot$, and a small number of evolved stars. These are reasonable heuristics for verifying that we successfully constructed a planet-blind occurrence survey, since a bias towards known giant planet hosts could manifest as a metal-rich sample (Fischer, Laughlin, et al., 2005), a particularly massive sample, or an excess of evolved stars (Johnson, Bowler, et al., 2010).

Since the CLS drew from the CPS RV catalog, our sample encompasses stars from several Keck-HIRES occurrence surveys, including Cumming et al., 2008 and Howard, Marcy, Johnson, et al., 2010. We refer the reader to Rosenthal, Fulton, et al., 2021 for the full star and planet catalog, as well as details of the planet search and completeness characterization.

Table 4.1: Small Planet Sample

Name	$M \sin i$ [$M_\oplus$]	a [AU]
HD 107148 b	$19.9^{+3.1}_{-3.1}$	$0.1407^{+0.0018}_{-0.0019}$
HD 115617 b	$16.1^{+1.1}_{-1.2}$	$0.2151^{+0.0028}_{-0.0029}$
HD 115617 c	$5.11^{+0.53}_{-0.51}$	$0.04956^{+0.00065}_{-0.00067}$
HD 11964A b	$24.4^{+2.0}_{-2.0}$	$0.2315^{+0.0021}_{-0.0022}$
HD 1326 b	$5.43^{+0.42}_{-0.42}$	$0.0732^{+0.00047}_{-0.00048}$
HD 141004 b	$13.6^{+1.5}_{-1.4}$	$0.1238^{+0.002}_{-0.002}$
HD 1461 b	$6.6^{+0.61}_{-0.56}$	$0.0636^{+0.00095}_{-0.00099}$
HD 1461 c	$7.07^{+0.88}_{-0.90}$	$0.1121^{+0.0017}_{-0.0017}$
HD 147379A b	$30.7^{+3.7}_{-3.8}$	$0.3315^{+0.0024}_{-0.0024}$
HD 156668 b	$5.03^{+0.42}_{-0.42}$	$0.05024^{+0.00051}_{-0.00052}$
HD 164922 b	$14.3^{+1.1}_{-1.1}$	$0.3411^{+0.0039}_{-0.0039}$
HD 164922 c	$10.53^{+0.98}_{-0.98}$	$0.2292^{+0.0026}_{-0.0027}$
HD 164922 d	$4.73^{+0.66}_{-0.66}$	$0.1023^{+0.0012}_{-0.0012}$
HD 168009 b	$9.5^{+1.2}_{-1.2}$	$0.1192^{+0.0017}_{-0.0018}$
HD 190360 b	$21.44^{+0.85}_{-0.84}$	$0.1294^{+0.0017}_{-0.0017}$
HD 192310 b	$14.3^{+2.0}_{-1.9}$	$0.3262^{+0.0036}_{-0.0037}$
HD 216520 b	$10.4^{+1.1}_{-1.2}$	$0.1954^{+0.0025}_{-0.0025}$
HD 219134 b	$16.41^{+1.00}_{-0.95}$	$0.2345^{+0.0027}_{-0.0027}$
HD 219134 c	$4.12^{+0.33}_{-0.34}$	$0.03838^{+0.00044}_{-0.00044}$
HD 219134 d	$7.73^{+0.73}_{-0.69}$	$0.1453^{+0.0017}_{-0.0016}$
HD 219134 e	$3.57^{+0.43}_{-0.45}$	$0.06466^{+0.00074}_{-0.00073}$
HD 285968 b	$9.1^{+1.4}_{-1.4}$	$0.06649^{+0.00043}_{-0.00043}$
HD 42618 b	$15.2^{+1.8}_{-1.8}$	$0.5337^{+0.0088}_{-0.0091}$
HD 45184 b	$11.9^{+1.3}_{-1.2}$	$0.0641^{+0.0011}_{-0.0011}$
HD 45184 c	$10.9^{+1.8}_{-1.8}$	$0.1095^{+0.0018}_{-0.0018}$
HD 69830 b	$10.26^{+0.69}_{-0.64}$	$0.0794^{+0.0012}_{-0.0012}$
HD 69830 c	$9.86^{+0.97}_{-0.94}$	$0.1882^{+0.0029}_{-0.0029}$
HD 69830 d	$14.1^{+1.7}_{-1.8}$	$0.645^{+0.01}_{-0.01}$
HD 75732 b	$9.37^{+0.43}_{-0.43}$	$0.01583^{+0.00024}_{-0.00024}$
HD 7924 b	$8.23^{+0.45}_{-0.44}$	$0.05595^{+0.00075}_{-0.00078}$
HD 7924 c	$8.83^{+0.63}_{-0.59}$	$0.1121^{+0.0015}_{-0.0016}$
HD 7924 d	$6.1^{+0.68}_{-0.65}$	$0.1532^{+0.0021}_{-0.0021}$
HD 90156 b	$11.8^{+2.0}_{-1.9}$	$0.2509^{+0.0036}_{-0.0037}$

Table 4.1: Small Planet Sample (Continued)

Name	$M \sin i$ [$M_\oplus$]	a [AU]
HD 95735 b	$18.0^{+2.9}_{-2.6}$	$3.1^{+0.13}_{-0.11}$
HD 97101 b	$10.2^{+1.3}_{-1.2}$	$0.2403^{+0.0017}_{-0.0017}$
HD 97658 b	$7.85^{+0.57}_{-0.55}$	$0.0805^{+0.0010}_{-0.0011}$
HD 99492 b	$26.7^{+1.9}_{-1.9}$	$0.1231^{+0.0014}_{-0.0015}$
GL 687 b	$17.6^{+1.5}_{-1.5}$	$0.1658^{+0.0012}_{-0.0012}$
HIP 74995 b	$16.22^{+0.62}_{-0.61}$	$0.04099^{+0.00042}_{-0.00044}$
HIP 74995 c	$5.06^{+0.69}_{-0.69}$	$0.07359^{+0.00076}_{-0.00078}$
HIP 57087 b	$21.22^{+0.70}_{-0.69}$	$0.02849^{+0.0002}_{-0.0002}$
GL 876 b	$5.86^{+0.50}_{-0.49}$	$0.02183^{+0.00018}_{-0.00019}$

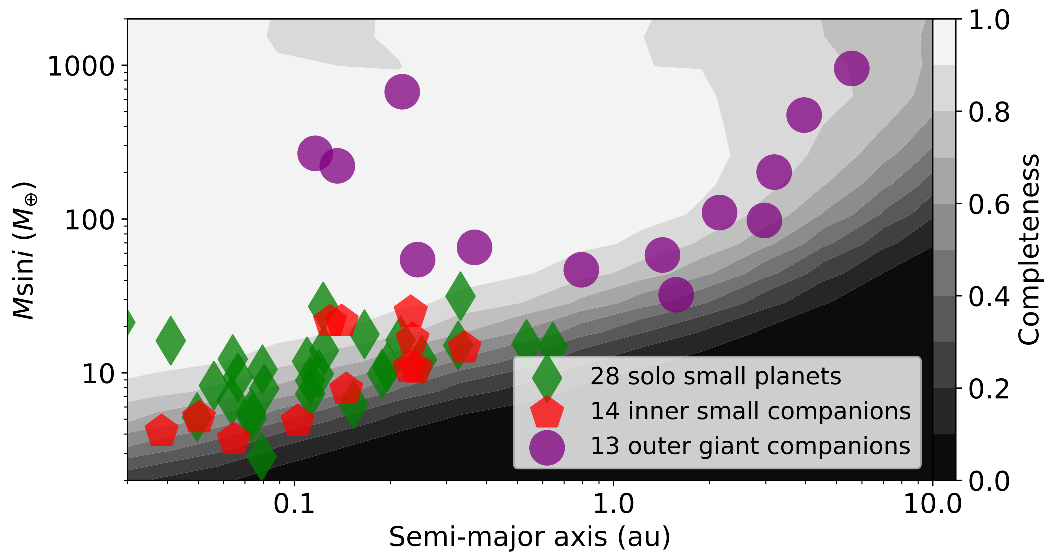

Figure 4.1: Minimum mass ($M \sin i$) versus semi-major axis values for small ($M \sin i < 30$ $M_\oplus$) planets, and their giant outer companions, in the CLS catalog. Diamonds are small planets without outer giants, pentagons are small planets with outer giants, and circles are outer giants. Contours are the completeness map for small planet hosts.

4.3 Methods

Occurrence model

The primary goal of this work is to measure planet occurrence, particularly of small close-in planets and cold gas giants. Many studies of RV or transit surveys use the intuitive occurrence measurement method known as "inverse detection efficiency" (Howard, Marcy, Bryson, et al., 2012; Petigura, Howard, et al., 2013). According to this procedure, one estimates occurrence in a region of parameter space by counting up the planets found in that region, with each planet weighted by the search completeness in that region. We measured the search completeness map of our survey by injecting many synthetic signals into each dataset, and computing the fraction of signals in a given region that are recovered by our search algorithm, `RVSearch`. Inverse detection efficiency as defined in Foreman-Mackey, Hogg, and Morton, 2014 is actually a specific case of a Poisson likelihood method, in which one models an observed planet catalog as the product of an underlying Poisson process and empirical completeness map.

Following the analysis in Fulton, Rosenthal, et al., 2021, we used the Poisson likelihood method to model the occurrence of planets. Given a population of observed planets with orbital and $M \sin i$ posteriors $\{\boldsymbol{\omega}\}$, and associated survey completeness map $Q(\boldsymbol{\omega})$, and assuming that our observed planet catalog is generated by a set of independent Poisson process draws, we can evaluate a Poisson likelihood for a given occurrence model $\Gamma(\boldsymbol{\omega}|\boldsymbol{\theta})$, where $\boldsymbol{\theta}$ is a vector parameterizing the rates of the Poisson process. The observed occurrence $\hat{\Gamma}(\boldsymbol{\omega}|\boldsymbol{\theta})$ of planets in our survey can be modeled as the product of the measured survey completeness and some underlying occurrence model,

$$\hat{\Gamma}(\boldsymbol{\omega}|\boldsymbol{\theta}) = Q(\boldsymbol{\omega})\Gamma(\boldsymbol{\omega}|\boldsymbol{\theta}). \tag{4.1}$$

The Poisson likelihood for an observed population of objects is

$$\mathcal{L} = e^{-\int \hat{\Gamma}(\boldsymbol{\omega}|\boldsymbol{\theta})\, d\boldsymbol{\omega}} \prod_{k=1}^{K} \hat{\Gamma}(\boldsymbol{\omega}_k|\boldsymbol{\theta}), \tag{4.2}$$

where K is the number of observed objects, and $\boldsymbol{\omega}_k$ is the kth planet's orbital parameter vector. The Poisson likelihood can be understood as the product of the probability of detecting an observed set of objects (the product term in Equation 2) and the probability of observing no additional objects in the considered parameter space (the integral over parameter space). Equations 1 and 2 serve as the foundation for our occurrence model, but do not take into account uncertainty in our measurements of planetary orbits and minimum masses. In order to do this, we use Markov Chain Monte Carlo methods to empirically sample the orbital posteriors of each system (Foreman-Mackey, Hogg, Lang, et al., 2013a; Fulton, Petigura, Blunt, et al., 2018; Rosenthal, Fulton, et al., 2021). We can hierarchically model the orbital posteriors of each planet in our catalog by summing our occurrence model over many posterior samples for each planet. The hierarchical Poisson likelihood is therefore approximated as

$$\mathcal{L} \approx e^{-\int \hat{\Gamma}(\boldsymbol{\omega}|\boldsymbol{\theta})\, d\boldsymbol{\omega}} \prod_{k=1}^{K} \frac{1}{N_k} \sum_{n=1}^{N_k} \frac{\hat{\Gamma}(\boldsymbol{\omega}_k^n|\boldsymbol{\theta})}{p(\boldsymbol{\omega}_k^n|\boldsymbol{\alpha})}, \tag{4.3}$$

where N_k is the number of posterior samples for the kth planet in our survey, and $\boldsymbol{\omega}_k^n$ is the nth sample of the kth planet's posterior. $p(\boldsymbol{\omega}|\boldsymbol{\alpha})$ is our prior on the individual planet posteriors. We placed uniform priors on $\ln(M \sin i)$ and $\ln(a)$. We used `emcee` to sample our hierarchical Poisson likelihood, and placed uniform priors on θ.

Approach to planet multiplicity

We want to evaluate the link between the presence of any inner small planets and the presence of any cold gas giants. Therefore, for all combinations of the presence or absence of these two planet types, we are interested in estimating the probability that a star hosts at least one planet. This quantity is distinct from the number of planets per star, both because many stars host more than one small planet (Howard, Marcy, Bryson, et al., 2012; Fang and Margot, 2012; He, Ford, Ragozzine, and Carrera, 2020) and because the probability of hosting at least one planet must be less than 1. We attempt to resolve this issue with two constraints on our model. First, we place a hard-bound prior on the integrated occurrence rate, so that it has an upper limit of one planet per star. Second, in the case of planetary systems that contain multiple detected planets in the class of interest, we only count the planet that was first detected by our search algorithm. We also report expected number of planets per star in Table 4.2, by including all companions in multi-planet systems.

The resulting estimate of the probability that a star hosts at least one planet depends on the search completeness in the mass and semi-major axis range of each individual planet. This biases our sample towards planets with greater RV semi-amplitudes, which tend to be closer-in and higher-mass. These planets are in higher-completeness regions and therefore will usually be detected first by iterative search algorithms. Figure 4.2 shows the observed multiplicity of the detected small planets in our sample. Note that this distribution is not corrected for search completeness, so it cannot be interpreted as the true underlying multiplicity distribution. Rather, it is showing how many multi-planet systems we detect with respect to systems where we only detect one small planet.

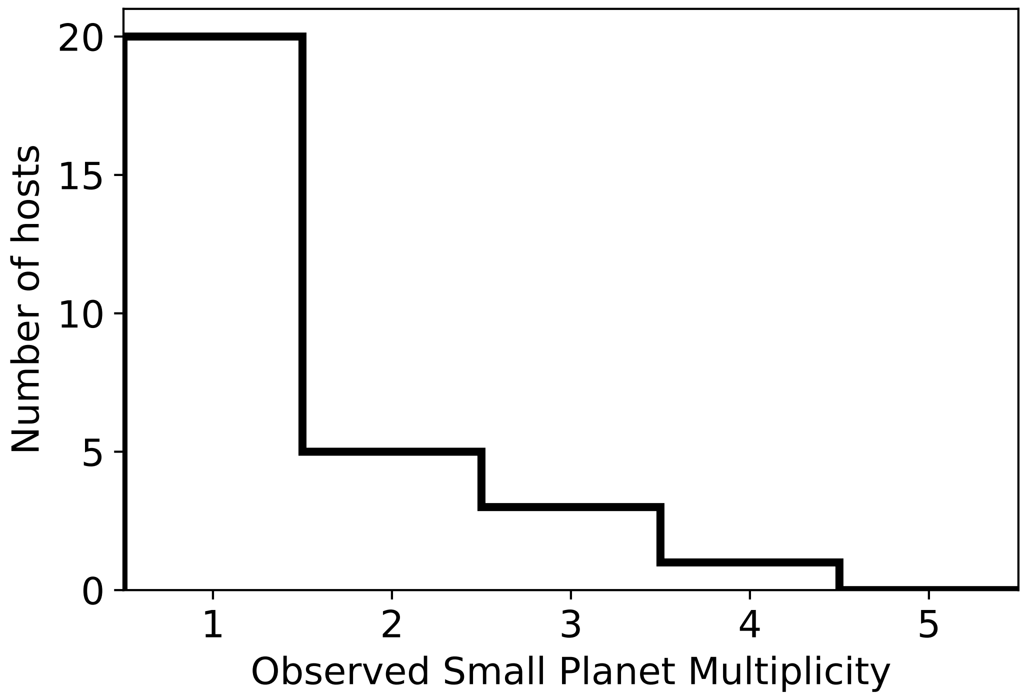

Figure 4.2: A histogram of observed small planet multiplicity in our sample. This is not corrected for search completeness, so it should only be interpreted as the multiplicity of detected planets, not as the underlying multiplicity distribution. There are 719 total stars in the CLS, around 29 of which we have detected small planets.

4.4 Results

Absolute and conditional occurrence rates

Using our occurrence methodology, we measured a set of distinct occurrence probabilities for the CLS sample. Specifically, we computed the absolute probability of hosting a small close-in planet, $P(I)$; the absolute probability of hosting a cold gas giant, $P(O)$; The probability of hosting a cold gas giant given the presence of a small close-in planet, $P(O|I)$; and the probability of hosting a small close-in planet given the presence of a cold gas giant, $P(I|O)$. In each case, we used our approach to multiplicity to link $P(\boldsymbol{\omega}|\boldsymbol{\theta})$ with $\Gamma(\boldsymbol{\omega}|\boldsymbol{\theta})$. We define the I range as 0.02–1 au and 2–30 $M_{\oplus}$. We define the O range in two ways: broadly, with 30–6000 $M_{\oplus}$and 0.23–10 au; and to only encompass Jupiter analogs as defined in Jones et al., 2016 and Bryan, Knutson, Lee, et al., 2019, with 3–7 au and 95–4130 $M_{\oplus}$. Figure 4.3 shows

$P(I|O)$ for the broad definition of giant planets, while Figure 4.4 shows $P(I|O)$ for Jupiter analogs. Table 4.2 reports all absolute and conditional probabilities for these populations, both for broad gas giants and Jupiter analogs. It shows that $P(I)$ and $P(I|O)$ are not significantly separated from each other, at least partially because the uncertainty in our measurement of $P(I|O)$ is high. In the following subsections, we compute the significance of the separation between two probability distributions as

$$S = \frac{|\bar{P}_2 - \bar{P}_1|}{\sqrt{\sigma_{P_2}^2 + \sigma_{P_1}^2}}, \tag{4.4}$$

where $\bar{P}$ is the mean of a distribution and σ_P^2 is its variance.

Condition	P(Condition)	$< N_P >$
Inner	$0.276^{+0.058}_{-0.048}$	$0.279^{+0.055}_{-0.053}$
Outer	$0.176^{+0.024}_{-0.019}$	$0.247^{+0.022}_{-0.023}$
Jupiter	$0.072^{+0.014}_{-0.013}$	$0.078^{+0.013}_{-0.014}$
Outer|Inner	$0.41^{+0.15}_{-0.13}$	$0.47^{+0.15}_{-0.12}$
Jupiter|Inner	$0.133^{+0.097}_{-0.063}$	$0.20^{+0.12}_{-0.08}$
Inner|Outer	$0.42^{+0.17}_{-0.13}$	$0.69^{+0.19}_{-0.19}$
Inner|Jupiter	$0.32^{+0.24}_{-0.16}$	$0.34^{+0.24}_{-0.17}$

Table 4.2: Absolute and conditional probabilities and number of planets per star for inner small planets, outer giants, and Jupiter analogs.

The impact of outer giants on inner small planet occurrence, and vice versa

Table 4.2 shows that $P(I) = 0.276^{+0.058}_{-0.048}$, whereas $P(I|O) = 0.42^{+0.17}_{-0.13}$. This implies that outer giant planets, according to our broad definition, enhance the occurrence of inner small planets with $\sim 1\sigma$ significance. Also, $P(O) = 0.176^{+0.024}_{-0.019}$, whereas $P(O|I) = 0.41^{+0.15}_{-0.13}$. This implies that inner small planets enhance the occurrence of outer giant planets with 1.65σ significance. This significance decreases when we narrow our outer companions to Jupiter analogs instead of a broad range of cold giants. In that case, $P(J|I)$ is only 0.85σ enhanced over $P(J)$, and $P(I|J)$ is not separated from $P(I)$. Additionally, whether we select a broad range of cold gas giants or a specific set of Jupiter analogs, our results rule out a 100% occurrence of small inner planets within 2–30 $M_\oplus$to outer gas giants.

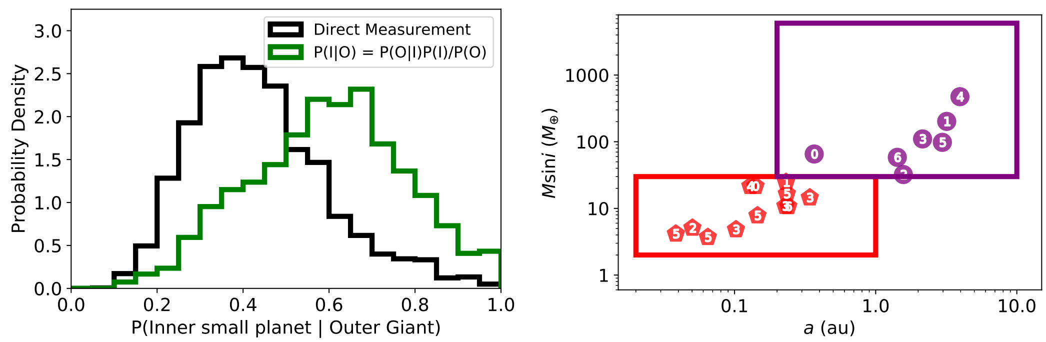

Figure 4.3: Left: Two measurements of the conditional occurrence of inner small planets given the presence of an outer gas giant. The black distribution is our direct measurement, while the green distribution uses Bayes Theorem to infer it from other measurements. Right: Our sample of planet pairs with small planets within the region of interest, with our inner small planet box outlined in red and our outer giant box outlined in purple. We assign a number to each planetary system and label individual planets accordingly.

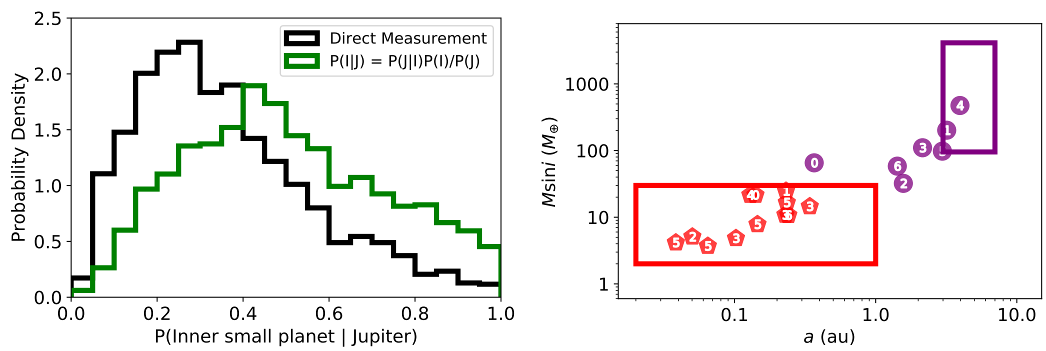

Figure 4.4: Same as Figure 4.3, but for Jupiter analogs within 3–7 au and 0.3–13 M_J instead of the broader giant population.

For the purposes of this work, we have assumed a small planet mass distribution that is uniform in $\ln(M)$. Assuming a uniform distribution in $\ln(a)$, this leads to a 25% recovery rate in our survey of small planets within 2–30 $M_\oplus$and 0.023–1 au, given our search completeness. Neil and Rogers, 2020 fit a joint mass-radius-period distribution to a sample of *Kepler* planets, and found a small planet mass distribution that is approximately log-normal, with mean $\mu_{\ln(M)} = 0.62$ and $\sigma_{\ln(M)} = 2.39$. Figure 4.5 plots and compares these two distributions. We find that assuming this log-normal distribution changes our average planet recovery rate within our small planet box from 25% to 18.3%, which would increase our corresponding

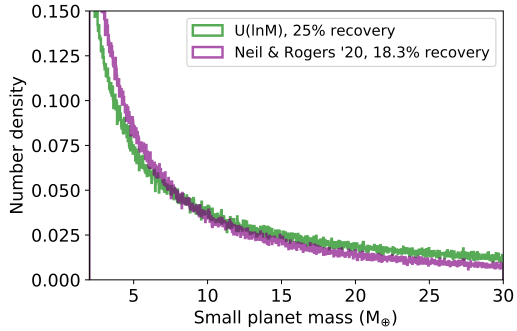

Figure 4.5: Analytical mass distributions for small planets used in this work (green) and from Neil and Rogers, 2020 (purple). This work assumes a uniform distribution in ln(M). Assuming a uniform distribution in ln(a), this leads to a 25% recovery rate in our survey of small planets within 2–30 $M_\oplus$and 0.023–1 au, given our search completeness. Neil and Rogers, 2020 fit a log-normal mixture model to a sample of *Kepler* planets and found a distinct small planet component, shown here. This model leads to an 18.3% recovery rate in our survey.

occurrence rate for inner super-Earths in systems with outer gas giants from 42% to roughly 58%. We conclude that our choice of mass distribution constitutes an additional source of uncertainty that is comparable to our measurement errors.

0.3–3 au giant suppression of inner small planets

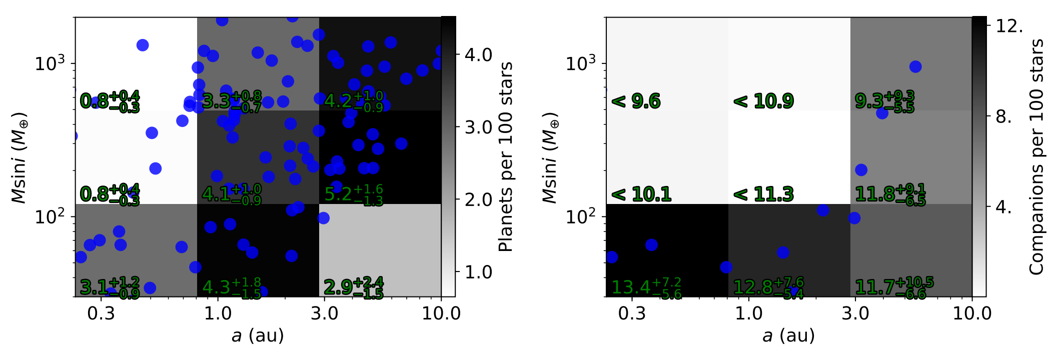

Figure 4.6: Left: Occurrence grid for the full CLS sample of 719 stars. Cell shade and number annotation reflect the median expected number of planets per 100 stars in each bin. Empty bins show an expected upper limit on occurrence as the 84.1th percentile on the occurrence rate posterior. Right: Same, but only for the 28 hosts of detected small planets with $M \sin i$< 30 $M_\oplus$and a < 1 AU. Note that this sample includes 55 Cnc's four cold giants, whereas our fractional analysis in Figures 4.3 and 4.4 excludes 55 Cnc, since its inner ultra-short period planet is undetectable by our automated search due to our period limits. The right-hand panel shows that there is an absence of warm gas giants in our sample of detected small planet hosts.

Figure 4.6, which shows occurrence grids of cold gas giants for our entire sample and for our sample of small planet hosts, provides tentative evidence that ′lukewarm′ giants within roughly 0.3–3 au may suppress small planet formation. The highest-completeness region of the small planet host parameter space, within 3 au and above ∼120 $M_{\oplus}$, is empty, whereas there are many detected giant planets in that region without detected small companions.

We can test the significance of our absence of lukewarm giants by referring to the broader distribution of gas giants, shown in the left panel of Figure 4.6, and calculating the probability of drawing 10 planets (our observed outer companions) from this distribution and finding 0 within the lukewarm Jupiter region. Normalizing the occurrence map shown in Figure 4.6, we find a 31.1% probability that a giant planet between 0.23–10 au will be found with $M \sin i > 120\ M_{\rm J}$ and $a < 3$ au, and 68.9% otherwise. We can simplify this test by using the binomial distribution to test the probability of drawing 0 out of 10 planets from a 31.1% lukewarm Jupiter probability, which simplifies to $0.689^{10} = 0.0241$. Thus, given our measured occurrence map for giant planets between 0.23–10 au, there is a 2.41% probability of drawing 10 planets from this population and seeing 0 lukewarm Jupiters more massive than 120 $M_{\rm J}$ and within 3 au. We drew this lukewarm boundary and performed this test after observing a paucity of lukewarm Jupiter companions, so it is possible that our result is biased by our sample. However, this definition of lukewarm Jupiter is physically motivated by mass and orbital separation, so it is more meaningful than a boundary arbitrarily chosen to exclude all planets.

The top panel of Figure 4.7 provides additional evidence that not all giant planets beyond 0.3 au host inner small planets, since the set of systems that host cold gas giants without detected small inner planets have non-zero sensitivity to said small planets, particularly within 0.1 au and above 7 $M_{\oplus}$. Independently, in a collection of 78 systems containing cold gas giants, we discover 5 systems with detected small planets in our small-planet range. Without a completeness correction, this yields a 6.41% probability of outer giants hosting inner small planets. Our completeness-aware methods yield 42%, a factor of ∼6.5x greater than this raw value. Conversely, the bottom panel of Figure 4.7 shows that the set of small inner planets without detected outer companions are sensitive to those companions with a completeness correction less than a factor of 1.5. Taken together, these results suggest that small planets are correlated with cold Jupiters, but potentially suppressed by warm Jupiters. The two exceptions in our sample are 55 Cnc, which hosts a very warm

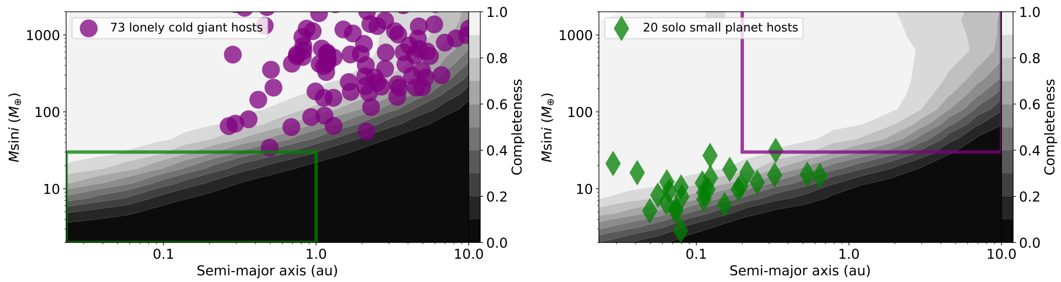

Figure 4.7: Top: cold giant planets without detected inner small companions in the CLS sample, with associated completeness contours. The contours show that the datasets associated with these systems are somewhat sensitive to planets within 0.5 au and 30 $M_\oplus$; we have outlined our small-planet parameter space in green for context. This means that we can say with some confidence that not all of these systems harbor undetected small planets. Bottom: Small planets without detected outer companions in the CLS sample, with associated completeness contours. We have outlined our outer giant parameter space in purple for context.

giant and an ultra-short period super-Earth, and GL 876, which hosts a 2-day sub-Neptune and a 2:1 resonant pair of super-Jupiters at 30 and 60 days. The fact that the two super-Jupiters are in an orbital resonance suggests that they likely formed farther out and then migrated inward (Yu and Tremaine, 2001), perhaps explaining how the inner sub-Neptune was able to form despite their presence.

Metallicity distributions

We used our sample to reproduce the previously derived result (Zhu and Wu, 2018; Bryan, Knutson, Lee, et al., 2019) that small planet hosts with outer gas giants are consistently more metal-rich with 97% significance than hosts of lonely small planets, as seen in Figure 4.8. This phenomenon agrees with the well-established correlation between metallicity and giant planet formation and, therefore, may be independent of the presence of small planets. However, since the CLS sample contains few systems with both outer giants and inner small planets, it is difficult to test the reverse effect and determine whether giant planet hosts with small planets have a different metallicity distribution than lonely giant planet hosts.

4.5 Discussion

Reconciling our results with other occurrence work and known systems

Bryan, Knutson, Lee, et al., 2019 found that $P(I|J) = 102^{+34}_{-51}\%$, while Zhu and Wu, 2018 found that $P(I|J) = 90 \pm 20\%$. Our measurements of $P(I|J) = 32^{+24}_{-16}\%$ and $P(I|O) = 42^{+17}_{-13}\%$ are consistent with the 2019 measurement, but are 2σ

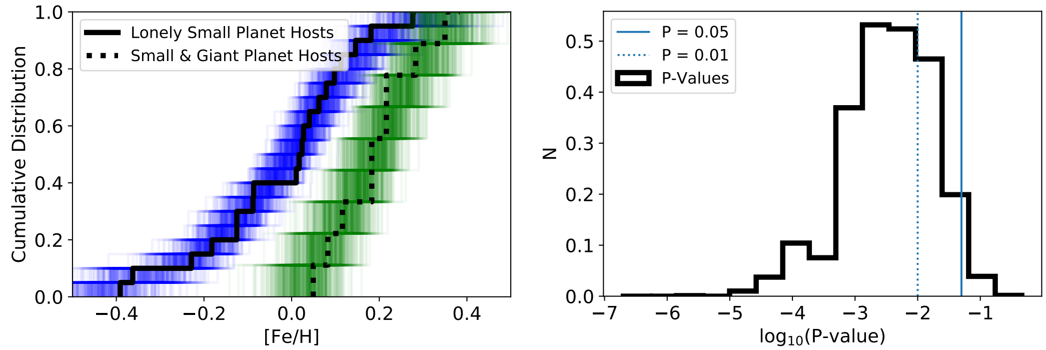

Figure 4.8: Left: Cumulative metallicity distributions for hosts of lonely small planets and hosts of both inner small and outer giant planets. Solid and dashed steps show distributions of median metallicity measurements, while transparent steps show metallicities drawn many times from Gaussian distributions, with means and standard deviations taken from measurement means and uncertainties. Right: A distribution of P-values from many Kolmogorov-Smirnov tests, performed on 5×10^5 drawn sets of metallicities for the two stellar host groups. 97% of the draws produce $P < 0.05$, and 73% of the draws produce $P < 0.01$. This implies that the underlying metallicity distributions of these two groups are distinct.

inconsistent with the 2018 measurement. Additionally, both of our measurements are more than 2.5σ separated from 100%.

Our finding that gas giants within a certain mass and semi-major axis range suppress the formation of inner small planets is highly conditioned on these mass and semi-major axis ranges. We limited this analysis to giant planets within 0.3–3 au. Conversely, Huang et al., 2016 found that half of all warm Jupiters have small planet companions by performing a similar analysis on the *Kepler* sample. They defined a warm Jupiter as a giant planet within 10–200 days. 200 days corresponds to a ~0.67 au orbit around a solar mass star, which is only beyond the inner limit of our 'lukewarm' range by about a factor of 2. This implies that our two results are not necessarily incompatible. Rather, they are drawn from mostly separate giant planet populations, which may have distinct formation or migration mechanisms.

Additionally, while the CLS does not contain 0.3–3 au giant companions to small planets, the *Kepler* sample contains several known systems that fit this description. For instance, the Kepler-167 system contains a 1 $M_J$giant at 1.9 au with three super-Earths, and the *Kepler*-1514 system contains a 5 $M_J$giant at 0.75 au with an inner 1.1 $R_\oplus$planet at 0.1 AU (Kipping et al., 2016; Dalba et al., 2021). Several other *Kepler* systems contain planets that satisfy or almost satisfy our criteria for small-and-giant pairs (Morton et al., 2016; Holczer et al., 2016), as well as non-*Kepler* systems (Bouchy et al., 2009; Stassun et al., 2017). We have not claimed that 0.3–3 au giants completely prevent the formation of small inner planets, only that these giants host inner small planets within 2–30 $M_\oplus$with a significantly smaller frequency than giant planets outside this range. Looking to the future, long-baseline RV follow-up of a very large sample of hosts of close-in small planets, such as a subsample of the *TESS* survey (Ricker et al., 2015), may uncover a larger number of outer giant companions. This would help clarify the precise distribution of these companions in mass and semi-major axis space.

Comparison between direct measurement and Bayesian inference of $P(I|O)$

Figure 4.3 shows that our direct estimate of $P(I|O)$ is lower than our indirect estimate using Bayes theorem, which calculates $P(I|O)$ as a function of $P(O|I)$, $P(O)$, and $P(I)$. These probabilities are likely mismatched because they assume uniform occurrence across giant planet parameter space, and Figure 4.6 shows that this is not the case. While our broad sample of giant planets fills $M \sin i$ and semi-major axis space, we found no outer companions to small planets among our warm Jupiters, as discussed in Subsection 4.4. This means that our population of outer giant companions and broader giant planet sample are distinct in parameter space, and that choosing a wide swath of $M \sin i$ and semi-major axis space for our Bayesian inference is not justified. This would also explain why our Jupiter analog comparison, shown in Figure 4.4, shows a closer match between a direct measurement and Bayesian inference. 3–7 au and 0.3–13 M_J is a narrower range of parameter space, and well separated from the warmer giants that appear to suppress small planet formation. This difference in giant classification could explain why the two posteriors more closely agree for the narrow definition of Jupiter analogs than for the broader definition that includes all cold gas giants.

The nature of cold giant companions

Figure 4.9 shows both all giant planets and outer giant companions to small planets in eccentricity, $M \sin i$, and semi-major axis space. The outer companions have an upper limit on eccentricity within 0.4, whereas the broader sample follows the beta distribution first described in Kipping, 2013. Figure 4.10 marginalizes over the occurrence distributions shown in Figure 4.6 to produce mass functions for these two populations within 0.23–10 au. This marginalization shows that outer companions are more frequently found at lower masses than the broader giant sample, with

$\sim 2\sigma$ significance. Figure 4.11 shows histograms of the maximum a posteriori eccentricities of all giant planets and outer giant companions to small planets. The broad sample has a moderate-to-high-eccentricity tail that is not shared by the outer companions. This makes intuitive sense, since eccentric giants are disruptive to the inner regions of a planetary system, and can disrupt the early-stage formation of small planets by sweeping through protoplanetary disks, or dynamically scatter small planets.

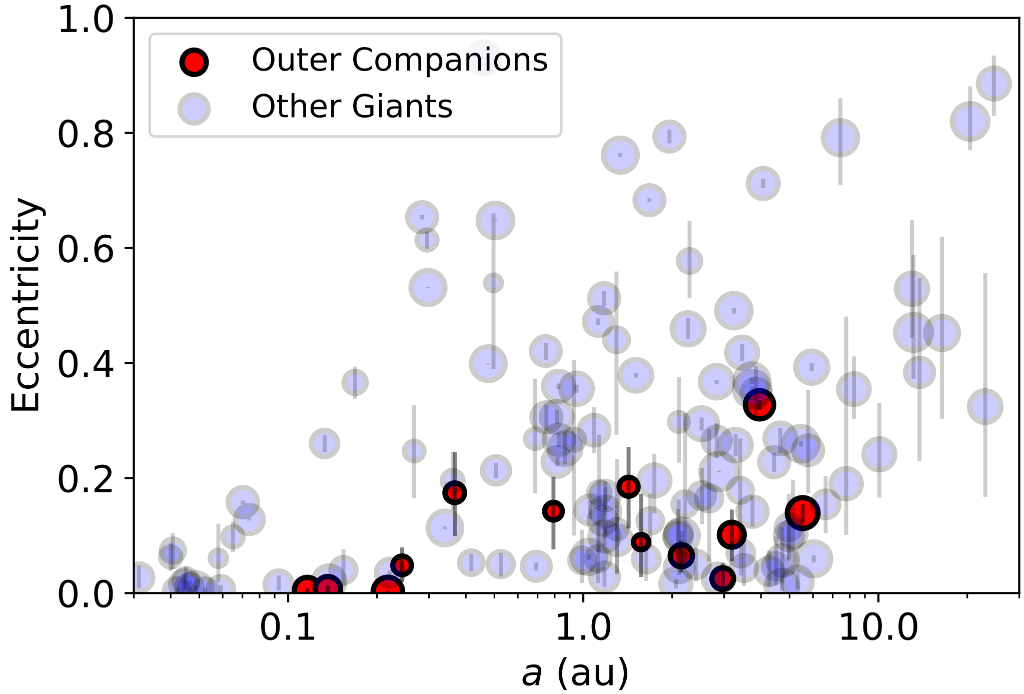

Figure 4.9: A comparison between outer giant companions and the rest of the CLS giant sample in semi-major axis and orbital eccentricity space. Circle size is proportional to $\log(M \sin i)$. Companions to inner small planets are less eccentric than the parent sample, and less massive.

An aside on our sample selection criteria

Our selection criteria differ from those of Bryan, Knutson, Lee, et al., 2019 in a number of ways. Our lowest mass planet has $M \sin i$ equal to 3.57 $M_\oplus$, below which we are almost entirely insensitive even to close-in planets. We therefore select a mass range of 2–30 $M_\oplus$, as opposed to 1–10 $M_\oplus$. We chose our upper limit on $M \sin i$ as an estimate of the mass threshold for runaway gas accretion (Lissauer et al., 2009). This limit also happens to correspond to a possible valley in the mass-radius distribution of planets, as seen in Neil and Rogers, 2020. In order to test whether

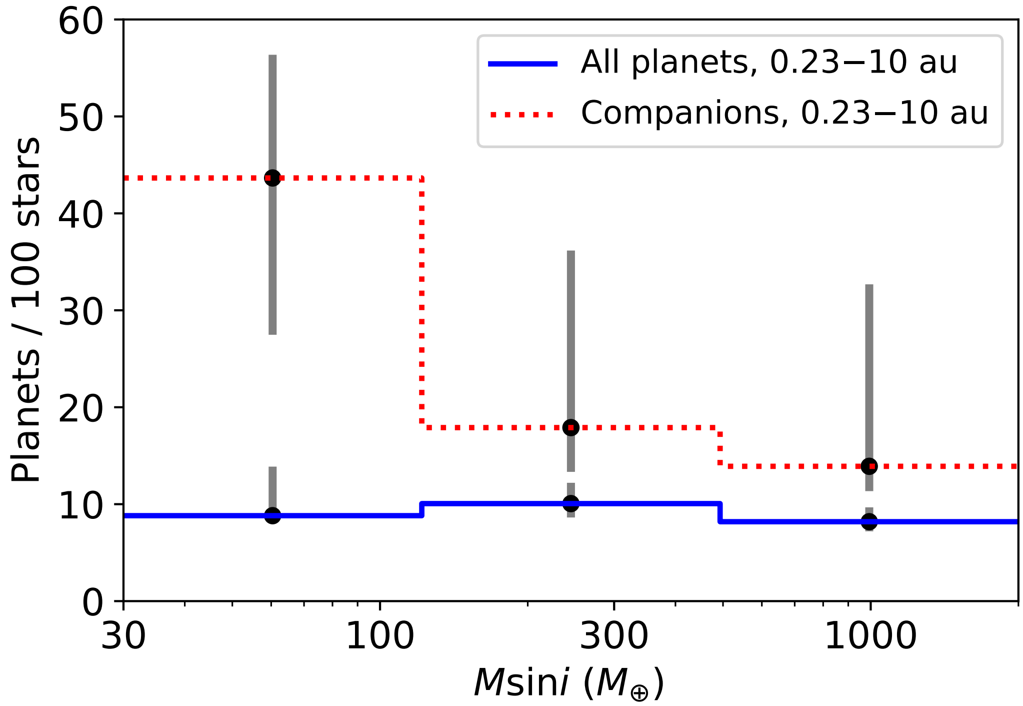

Figure 4.10: Mass functions for all planets and outer companions within 0.23–10 au, produced by marginalizing the occurrence grids in Figure 4.6 along their orbital separation axes. Steps and dots show maximum a posteriori values; vertical bars show 15.9–84.1% confidence intervals.

this difference will significantly challenge our comparison to the prior results, we recompute our measurements with tighter limits on small planet parameter space, moving to 2–20 $M_\oplus$and 0.023–0.5 au. We show these results in Table 4.3. $P(I|O)$ is consistent for both definitions of an inner small planet; $P(O|I)$ is $\sim 1\sigma$ distinct. We cannot compare Jupiter analog conditional probabilities because the narrower sample of small planets does not include any companions to Jupiter analogs. These results imply that our choice of $M \sin i$ and a limits for small inner planets do not significantly impact our comparisons to studies with different definitions of small planets. We also explored the impact of changing our lower $M \sin i$ limit from 2 $M_\oplus$to 3 $M_\oplus$, since there are no companion small planets in our sample that are less massive than 3 $M_\oplus$. These changes decreased $P(I|O)$ and $P(I)$ by less than 1σ and 1.5σ respectively, as seen in Table 4.3.

Condition	(1 au, 2–30 $M_\oplus$)	(0.5 au, 2–20 $M_\oplus$)	(0.5 au, 3–30 $M_\oplus$)
Inner	$0.276^{+0.058}_{-0.048}$	$0.281^{+0.066}_{-0.051}$	$0.191^{+0.036}_{-0.035}$

Outer|Inner	$0.41^{+0.15}_{-0.13}$	$0.29^{+0.14}_{-0.11}$	$0.43^{+0.17}_{-0.13}$
Inner|Outer	$0.42^{+0.17}_{-0.13}$	$0.46^{+0.20}_{-0.16}$	$0.28^{+0.12}_{-0.09}$
Jupiter|Inner	$0.133^{+0.097}_{-0.063}$	No detections	$0.21^{+0.12}_{-0.09}$
Inner|Jupiter	$0.32^{+0.24}_{-0.16}$	No detections	$0.25^{+0.17}_{-0.13}$
Outer	$0.176^{+0.024}_{-0.019}$		
Jupiter	$0.072^{+0.014}_{-0.013}$		

Table 4.3: An expanded version of Table 4.2's set of probabilities, with columns for small planets within 0.023–0.5 au and either 2–20 $M_\oplus$ or 3–30 $M_\oplus$.

Additionally, we did not search for planets with orbital periods less than one day, since this would produce alias issues in our automated search pipeline. This leads to a complication regarding 55 Cnc, which hosts a super-Earth with an orbital period of 0.74 days. This planet is the only previously known USP in our sample, and this system is one of the few that we initialized with known planets in our search, including a Keplerian orbit for the USP in order to properly model our RV data. This system also stands out from the rest of our sample in a number of other ways, such as hosting both a hot Jupiter and multiple outer, less massive giants. Since our blind search does not extend below 1 day, we should in principle limit our small-planet sample to planets beyond 0.02 au, which corresponds to just over a 1-day orbit around a G dwarf. This excludes 55 Cnc and its giant planets from our Bayesian estimates of inner and outer companion probability and leaves only 2 planets within the Jupiter analog box instead of 3. We opted to exclude 55 Cnc from our conditional probability analysis for the sake of consistency, but left it in our outer giant occurrence grids shown in Figure 4.6. Redoing the analysis with an inner limit of 0.015 au, so that 55 Cnc is included, we find that the probability of hosting a close-in small planet given the presence of an outer Jupiter analog is $0.39^{+0.21}_{-0.16}$, as opposed to $0.32^{+0.24}_{-0.16}$ without 55 Cnc. Likewise, when including 55 Cnc, the probability of hosting a close-in small planet given the presence of a cold gas giant in broader parameter space (0.23–10 au, 30–6000 $M_\oplus$) is $0.42^{+0.17}_{-0.13}$, as opposed to $0.42^{+0.17}_{-0.12}$ without 55 Cnc, i.e., nearly identical.

Alongside 55 Cnc, GJ 876 is the one other system in our sample that hosts both a detected small planet and warm gas giants. This system hosts a small planet on a 2-day orbit and two giant planets in a 2:1 resonance at 30 and 60 days. We propose that this system is the exception that supports our theory of warm Jupiters

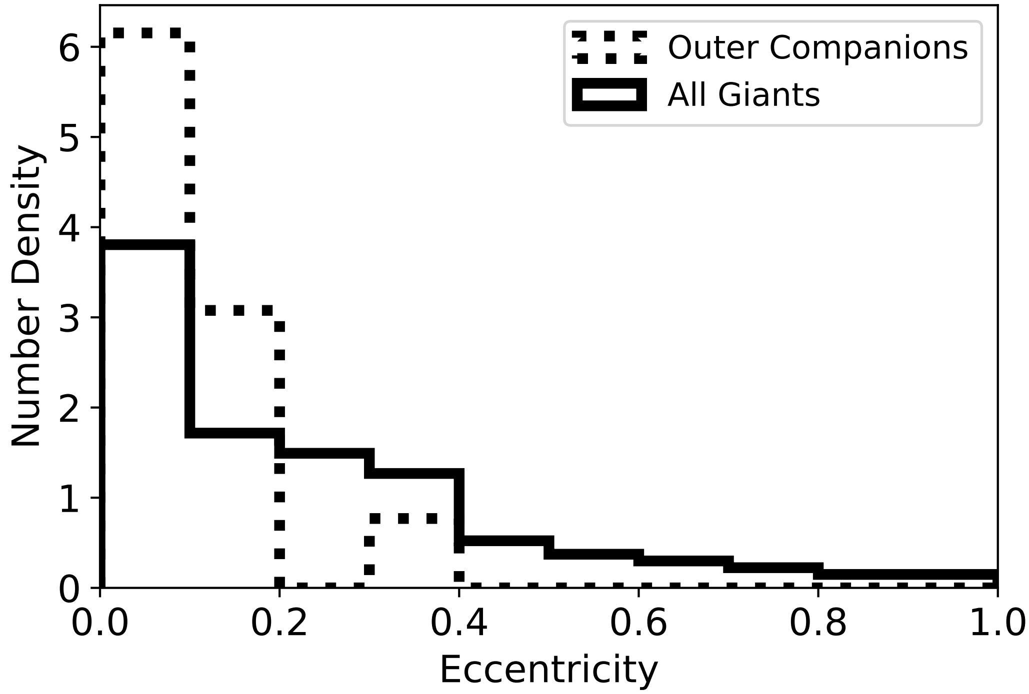

Figure 4.11: Observed eccentricity distributions for all giant planets and outer giant companions. This plot shows histograms of the maximum a posteriori eccentricities of our sample.

suppressing inner small planet formation, since this resonant pair may have migrated inward from beyond 1 au (Yu and Tremaine, 2001; Batygin, 2015; Nelson et al., 2016).

An aside on multiplicity bias

We investigate whether our estimate of the probability that a star hosts at least one planet ($P(1+)$) in a given parameter space is systematically biased in cases of high planet multiplicity. We estimate the possible magnitude of this effect using a Monte Carlo experiment to recover the true value of $P(1+)$ given a toy model for planet multiplicity and simulated observed population. First, we choose an underlying probability of hosting at least one planet p, as well as a simple multiplicity distribution f_n given the presence of at least one planet, capped at 3 planets.

Then, we perform a single step of our Monte Carlo experiment by 'creating' 81 stars, the size of our outer giant O host sample. Each star has a probability p of hosting any inner small planets, and probability f_n of hosting n such planets, if it hosts any such planets at all. We sample these planets from uniform $\ln(a)$ and $\ln(M)$ distributions in the desired parameter space, in our case our small planet definition. We then determine how many planets we detect around each star by only keeping the planets with $\ln(a)$ and $\ln(M)$ pairs that have search completeness higher than a random number drawn from $U(0, 1)$. We randomly select a new completeness contour for each generated system.

Once we have our population of observed planets, we run our Poisson likelihood model on the first-observed planets to generate our estimate of the probability that a star hosts at least one planet in a given parameter space, or $P(1+)$. Figure 4.12 shows our results for $p = 0.3$ and three different choices of multiplicity, including the distribution of detected multis shown in Figure 4.2. When we assume one small planet per host, our Poisson model accurately retrieves the underlying probability of hosting at least one planet. With increasing average multiplicity, our model marginally overestimates $P(1+)$ by a small but increasing factor. This is because increasing the expected intrinsic number of planets per host increases the probability of detecting at least one planet around a true host.

Correcting for this bias would require confident knowledge of the multiplicity distribution of small planets, which is currently contested even after much work with *Kepler* (e.g., Zhu and Wu, 2018; He, Ford, and Ragozzine, 2019). Resolving this issue would require a much larger and more complete small planet sample than the one available through the CLS. Additionally, this bias analysis is specific to our Poisson likelihood methods, which are better suited to correcting for survey incompleteness than binomial or Bernouilli estimates. Hopefully, future precise RV exoplanet surveys will produce more rigorous small planet multiplicity measurements, and create opportunities to better understand bias in occurrence rates due to multiplicity.

Implications for planet formation

One key takeaway from our analysis is that lukewarm Jupiters may either suppress the formation or migration of small inner planets or destabilize the orbits of inner super-Earths. The first conjecture fits with theoretical work (Kley and Nelson, 2012; Moriarty and Fischer, 2015) that shows how gas giants that are sufficiently massive or close to their host stars can create gaps in the protoplanetary disk and prevent the inward flow of solids beyond their orbits. This cuts off the supply of material during the critical timescales of pebble accretion, thus depriving rocky cores of the fuel

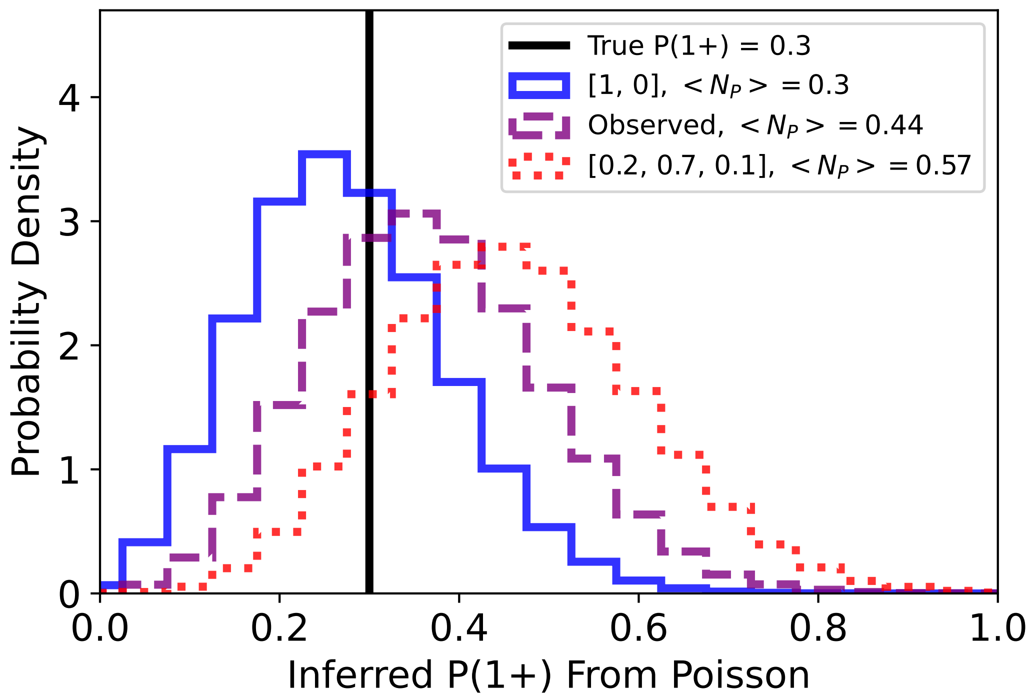

Figure 4.12: The results of three Monte Carlo experiments detailed in Subsection 4.5. We estimate how planet multiplicity biases our estimates of the probability that a star hosts at least one small planet, or $P(1+)$. The black vertical line is the true value of $P(1+)$ for all three experiments. The legend shows the multiplicity distribution and resulting average number of planets per host $< N_P >$ for all three experiments. In the one-per-host case, the Poisson model accurately recovers $P(1+)$, while cases with higher $< N_P >$ increasingly overestimate the true value.

needed to grow into super-Earths or larger planets (Chachan, Dalba, et al., 2022). Alternatively, warm or lukewarm gas giants may excite the eccentricities of nascent inner small planets and destabilize their orbits into ejection or accretion (Schlecker et al., 2020). Both of these explanations imply that a cold gas giant beyond 5 au such as Jupiter is not detrimental to interior small planet formation, but a Jupiter-mass giant within 0.3–3 au may be.

These ideas do not have to clash with the known coexistence of small planets and warm gas giants. Huang et al., 2016 found that warm Jupiters (10–200 days) with close companions are substantially more common than hot Jupiters (< 10 days) with close companions. It is possible that there is a warm Jupiter pile-up due to migration, which takes cold gas giants all the way through the region where we see a dearth of warmer giant companions to small inner planets.

4.6 Conclusion

We explored the relationship between small, close-in planets and outer giants by computing absolute and conditional probabilities for these two populations. We found that $42^{+17}_{-13}\%$ of stars that host a giant planet within 0.23–10 au also host an inner small planet between 2 and 30 $M_{\oplus}$, and that $32^{+24}_{-16}\%$ of stars that host a Jupiter analog (3–7 au, 0.3–13 M_{J}) also host an inner small planet between 2 and 30 $M_{\oplus}$. These probabilities are $\sim 1\sigma$ separated from the absolute probability of hosting a small close-in planet, implying an inconclusive effect of outer gas giants on the occurrence of small, close-in companions. On the other hand, the probability of hosting an outer gas giant given the presence of a small planet is 1.65σ enhanced over the absolute probability of hosting an outer gas giant. We also confirmed the known result that stars with both small, close-in planets and cold giants tend to be more metal-rich than stars with only small planets. Additionally, we used Monte Carlo simulations to estimate how small planet multiplicity might bias Poisson estimates of the probability that a star hosts at least one small planet. We found that multiplicity may result in overestimating this probability, but that assumptions in our Poisson model may reduce the magnitude of this bias.

All code used in this chapter is available at `github.com/California-Planet-Search/rvsearch` and `github.com/leerosenthalj/CLSIII`. This research makes use of GNU Parallel (Tange, 2011). We made use of the following publicly available Python modules: `astropy` (Astropy Collaboration et al., 2013), `matplotlib` (Hunter, 2007), `numpy/scipy` (Walt et al., 2011), `pandas` (McKinney, 2010), `emcee` (Foreman-Mackey, Hogg, Lang, et al., 2013a), and `RadVel` (Fulton, Petigura, Blunt, et al., 2018).

Chapter 5

LONELY, POOR, AND ECCENTRIC: A COMPARISON BETWEEN SOLITARY AND NEIGHBORLY GAS GIANTS

5.1 Introduction

Giant planets are central players in planetary systems. Their dynamics affect the presence of small inner planets (Zhu and Wu, 2018; Bryan, Knutson, Lee, et al., 2019; Rosenthal, Knutson, et al., 2021), and can reveal key information about the formation and dynamical history of a system. We can also better understand the impact that giant planets have on each other by studying their eccentricity distributions, as planets with non-zero eccentricities must have reached their states either by interacting with protoplanetary disks, or by undergoing dynamical interactions with other planets or stellar companions (Dawson and Murray-Clay, 2013; Petrovich and Tremaine, 2016). While secular interactions between planets and planet-disk effects can account for giants excited to moderate eccentricities ($0.2 \leq e \leq 0.6$) (Dawson and Johnson, 2018), it is more difficult to use them as an explanation for the highest-eccentricity giants that have been discovered to date, which extend to $e > 0.9$ (Blunt et al., 2019).

Prior work has shown that highly eccentric giants can reach their states due to strong gravitational scattering events, possibly with other giants (Chatterjee et al., 2008). This would imply that giant multiplicity is a key factor in understanding the dynamical evolution and final states of planetary systems. Perhaps systems in which we currently only see one eccentric giant began with multiple giant planets that scattered each other, one into a bound eccentric orbit and one into engulfment or an unbound trajectory. We can explore this theory by comparing the eccentricity distributions of systems with one observed giant planet and multiple observed giant planets, as well as their semi-major axis occurrence distributions and host star properties. Recent theoretical work has provided giant planet hypotheses that are easily testable with radial velocity (RV) surveys. For instance, Jackson et al., 2021 used a synthetic population to show that if an RV search for outer giant companions to warm Jupiters (10–200 days) finds that a substantial fraction of these warm giants have outer companions, then we should seriously consider a formation scenario in which secular interactions cause cold giants to migrate inward and become stable warm Jupiters.

The California Legacy Survey (CLS, Rosenthal, Fulton, et al., 2021) is well suited for a comparative study of lonely and neighborly giant planets. As a three-decade long blind RV survey, it produced a sample that is appropriate for a variety of occurrence measurements, and contains over a hundred giant ($M \sin i > 0.1 M_J$) planets, in both single and multiple configurations. In this paper, we leverage this survey to compare and contrast lonely giants and multi-giant systems. In Section 2, we review the star and planet catalog of the CLS. In Section 3, we describe our methods for computing planet occurrence, hierarchically modeling eccentricity distributions, and comparing the host star properties of distinct planetary samples. In Section 4, we present our results. In Section 5, we discuss our findings and their context.

5.2 Survey Review

The California Legacy Survey (Rosenthal, Fulton, et al., 2021) is a sample of RV-observed FGKM stars and their associated planets, created in order to provide a stellar and planetary catalog for occurrence studies. We approximated a quantifiably complete survey by selecting HIRES-observed stars that were originally chosen without bias towards a higher or lower than average likelihood of hosting planets. That left us with 719 stars. We used an iterative periodogram method to search for planet candidates, and performed uniform vetting to identify false positives.

This left us with 178 planets in our sample, including 134 planets with $M \sin i > 0.1 M_J$. 65 of these gas giants lack detected companions ('lonely giants'), and 69 belong to 31 multiple-giant systems. Figure 5.1 shows our giant catalog in eccentricity and semi-major axis space, split between giant singles and giant multis in one panel, and color-coded by host star metallicity in another panel. Figure 5.2 shows the catalog in eccentricity and metallicity space.

We use $0.1 M_J$ or 30 $M_\oplus$ as our cutoff for giant planets because this is roughly the minimum mass at which runaway gas accretion can begin (Lissauer et al., 2009). This is also near the threshold for planets to have compositions dominated by gaseous envelopes (Lee, 2019; Thorngren et al., 2016). These threshold bases set a more liberal definition of giant planet than one motivated by dynamical influence on planet formation and migration, but is more relevant to a planet's growth history and captures sub-Saturn planets that can still have an impact on smaller planets. Furthermore, our sample is reasonably insensitive to our definition of this threshold.

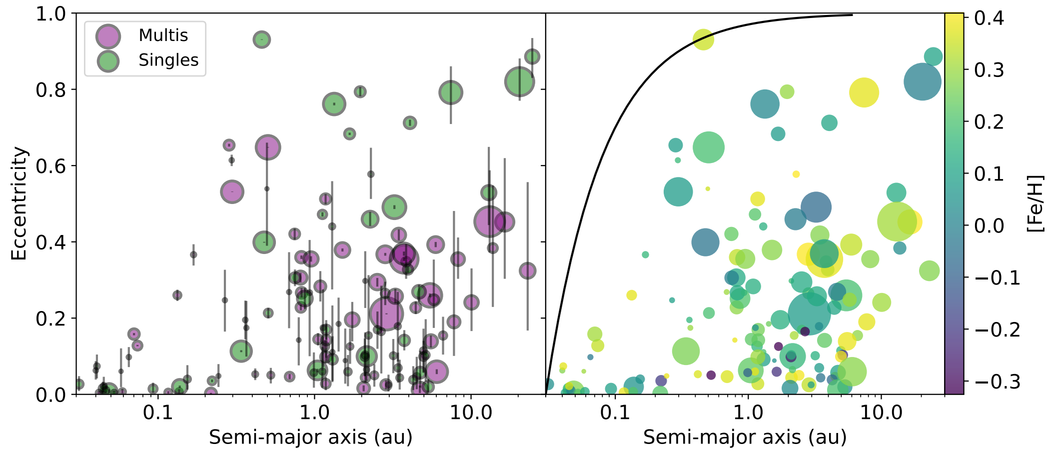

Figure 5.1: Left: Eccentricity with respect to semi-major axis for single giant planets and multiple giant planets. Vertical bars show 15.9–84.1% confidence intervals. Circle radius is proportional to $M \sin i$. Right: Same as left, but color-coded by host star metallicity. Curve shows estimate of tidal disruption limit assuming that periastron at 0.03 au leads to disruption (Dawson and Johnson, 2018).

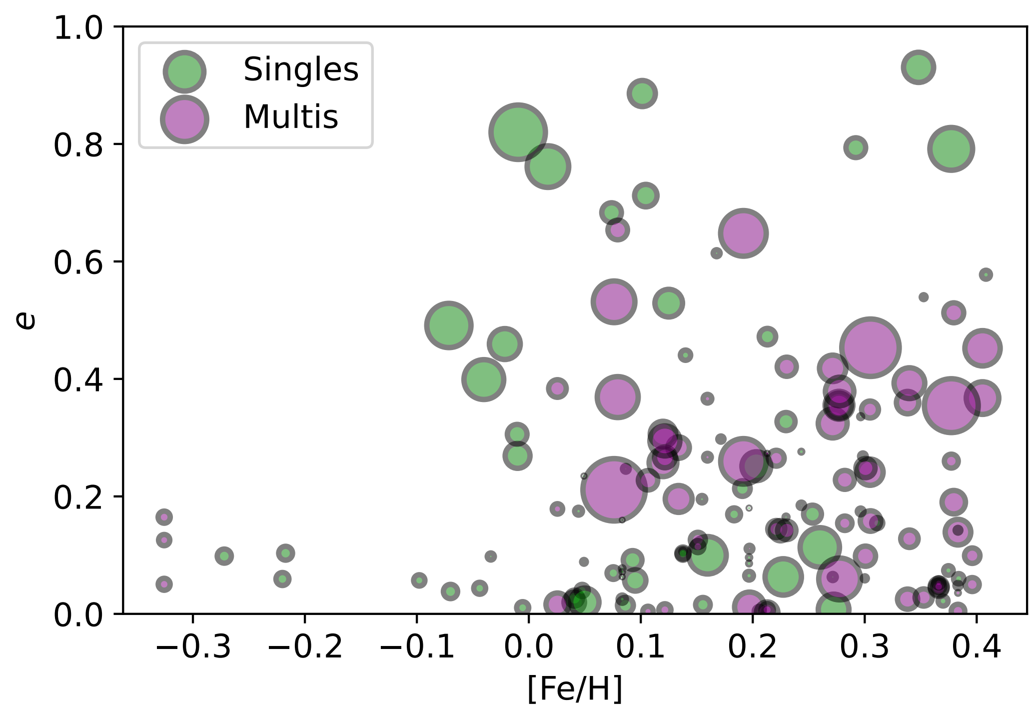

Figure 5.2: Planet eccentricity versus host star metallicity for single and multiple giant systems.

When we lower the threshold from 30 $M_\oplus$ to 20 $M_\oplus$ the number of single giant systems only decreases from 65 to 64, and the number of multi-giant systems only increases from 31 to 34.

5.3 Methods

Occurrence model

This work measures the occurrence of single and multiple giant planets as a function of multiple orbital properties. Many studies of RV or transit surveys use the intuitive occurrence measurement method known as "inverse detection efficiency" (Howard, Marcy, Bryson, et al., 2012; Petigura, Howard, et al., 2013). According to this procedure, one estimates occurrence in a region of parameter space by counting up the planets found in that region, with each planet weighted by the search completeness in that region. One can measure the search completeness map of a survey by injecting many synthetic signals into each dataset, and computing the fraction of signals in a given region that are recovered by the search algorithm in use. Inverse detection efficiency is actually a specific case of a Poisson likelihood method, in which one models an observed planet catalog as the product of an underlying Poisson process and empirical completeness map (Foreman-Mackey, Hogg, and Morton, 2014).

In this work, as in previous papers in the CLS (Fulton, Rosenthal, et al., 2021; Rosenthal, Knutson, et al., 2021), we used the Poisson likelihood method to model the occurrence of giant planets, taking measurement uncertainty into account. Given an observed population of observed planets with orbital and $M \sin i$ posteriors $\{\boldsymbol{\omega}\}$, and associated survey completeness map $Q(\boldsymbol{\omega})$, and assuming that our observed planet catalog is generated by a set of independent Poisson process draws, we can evaluate a Poisson likelihood for a given occurrence model $\Gamma(\boldsymbol{\omega}|\boldsymbol{\theta})$, where $\boldsymbol{\theta}$ is a vector of model parameters. The observed occurrence $\hat{\Gamma}(\boldsymbol{\omega}|\boldsymbol{\theta})$ of planets in our survey can be modeled as the product of the measured survey completeness and some underlying occurrence model,

$$\hat{\Gamma}(\boldsymbol{\omega}|\boldsymbol{\theta}) = Q(\boldsymbol{\omega})\Gamma(\boldsymbol{\omega}|\boldsymbol{\theta}). \tag{5.1}$$

The Poisson likelihood for an observed population of objects is

$$\mathcal{L} = e^{-\int \hat{\Gamma}(\boldsymbol{\omega}|\boldsymbol{\theta})\, d\boldsymbol{\omega}} \prod_{k=1}^{K} \hat{\Gamma}(\boldsymbol{\omega}_k|\boldsymbol{\theta}), \tag{5.2}$$

where K is the number of observed objects, and $\boldsymbol{\omega}_k$ is the kth planet's orbital parameter vector. The Poisson likelihood can be understood as the product of the probability of detecting an observed set of objects (the product term in Equation 2), and the probability of observing no additional objects in the considered parameter space (the integral over parameter space). Another way to understand this model is that it represents a Poisson process, with some probability of generating planets with a given mass and semi-major axis rate density.

Equations 1 and 2 serve as the foundation for our occurrence model, but do not take into account uncertainty in our measurements of planetary orbits and minimum masses. In order to do this, we use Markov Chain Monte Carlo methods to empirically sample the orbital posteriors of each system (Rosenthal, Fulton, et al., 2021). We can hierarchically model the orbital posteriors of each planet in our catalog by summing our occurrence model over many posterior samples for each planet. The hierarchical Poisson likelihood is therefore approximated as

$$\mathcal{L} \approx e^{-\int \hat{\Gamma}(\boldsymbol{\omega}|\boldsymbol{\theta})\, d\boldsymbol{\omega}} \prod_{k=1}^{K} \frac{1}{N_k} \sum_{n=1}^{N_k} \frac{\hat{\Gamma}(\boldsymbol{\omega}_k^n|\boldsymbol{\theta})}{p(\boldsymbol{\omega}_k^n|\boldsymbol{\alpha})}, \tag{5.3}$$

where N_k is the number of posterior samples for the kth planet in our survey, and $\boldsymbol{\omega}_k^n$ is the nth sample of the kth planet's posterior. $p(\boldsymbol{\omega}|\boldsymbol{\alpha})$ is our prior on the individual planet posteriors. We placed uniform priors on $\ln(M \sin i)$ and $\ln(a)$. We used `emcee` (Foreman-Mackey, Hogg, Lang, et al., 2013a) to sample our hierarchical Poisson likelihood.

Application to eccentricity

As done in prior eccentricity studies (Kipping, 2013; Van Eylen et al., 2019; Bowler, Blunt, et al., 2020), we model the eccentricity distribution of a population of exoplanets with the Beta distribution, because it is [0, 1] bound, flexible in its shape, and has only two model parameters. The Beta distribution follows the probability density function

$$p(x|\alpha,\beta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1}, \tag{5.4}$$

where $\Gamma(x)$ is the Gamma function, rather than the occurrence rate in Equation 3. We only consider planets with $M \sin i \geq 0.1\ M_{\mathrm{J}}$. In other words, we are only considering the eccentricity distributions of planets with masses greater than a third of Saturn's mass.

When we apply our hierarchical modeling framework to the eccentricity posteriors of our giant planets, we produce the likelihood

$$\mathcal{L}(\boldsymbol{e}|\boldsymbol{\theta}) \propto \prod_{k=1}^{K} \frac{1}{N} \sum_{n=1}^{N} p(e_k^n|\boldsymbol{\theta}), \tag{5.5}$$

where θ is the set of model parameters, K is the number of planets, N is the number of samples drawn from each planet's posterior, and e_k^n is the eccentricity of the nth sample drawn from the kth planet's posterior. The normalization term in Equation (3) disappears because Equation (4) is a normalized probability density function.

Empirical distribution comparisons

In order to compare the host-star metallicities and other properties of distinct groups of planetary systems, we use simple statistical tests to compare the cumulative distribution functions, otherwise known as empirical distributions, of these groups. Specifically, we use the Kolmogorov-Smirnov test, a statistical method for computing the probability that two empirical distributions were drawn from the same underlying distribution.

5.4 Results

Single and multiple giants have distinct eccentricity distributions

Figures 5.3 and 5.4 show our results when we use the hierarchical model described in Section 5.3 to characterize the distinct eccentricity distributions of single and multiple giant planets. While we define multiplicity as the number of giant planets at any semi-major axis, our joint fits only include planets beyond 0.3 au, since giants within that distance have short tidal circularization timescales (Ogilvie, 2014). Both the beta distribution models and the raw posteriors show that single giant planets have a pile-up of circular orbits and a long tail that extends to $e \leq 1$, while the multiple giant planets have a more extended range of moderate eccentricity and a sharp cutoff around $e \sim 0.7$. Hints of this behavior were originally visible in in Figure 13 of Wright, Upadhyay, et al., 2009, which compared the eccentricity distributions of all single and multi-planet systems discovered to date and found a long single-planet tail not shared by the multis. Bryan, Knutson, Howard, et al., 2016 reported a similar result with a larger planet sample.

One possible unifying explanation is that neighborly giants have some probability of experiencing secular interactions or scattering events, whereas lonely giants have no clear mechanisms for high-eccentricity excitation except for strong disk interactions or Kozai-Lidov interactions with a wide stellar binary companion. This would imply that the observed population of more circular lonely giants have not experienced companion-driven excitation, while the moderately eccentric multis and highly eccentric singles represent different outcomes from planet-planet interactions. Eccentric giants that appear to be lonely now may have been born with nearby companions and scattered them into ejection or tidal engulfment, while systems with two observed, moderately eccentric giants experienced dynamical interactions that were strong enough to excite but not eject or cause engulfment. We can test this theory by comparing the host metallicity distributions of these subgroups and checking whether the eccentric singles have more in common with circular singles or the multis.

Hosts of multiple giants are metal-richer than lonely giant hosts

Figure 5.5 compares the host star metallicity distributions of systems with only one observed giant planet and multiple observed giant planets. Hosts of multiple giant planets are distinctly more metal rich than hosts of only one detected giant planet, with 92.2% significance, that being the fraction of Kolmogorov-Smirnov tests performed on many samples from our distributions that fall within a p-value $p \leq 0.05$. This finding is in agreement with the trend first observed in Fischer, Laughlin, et al., 2005. We split our giant host sample along its median metallicity and found that the search completeness contours for the two subsamples are nearly identical, as seen in Figure 5.6. This provides confidence that the difference in these two distributions is not caused by more extensive RV observations of our most metal-rich host stars.

We can interpret this phenomenon in a number of ways. One possible explanation is that the probability of more than one giant planet forming around a star increases with increasing host metallicity. This is consistent with expectations from the core accretion theory of giant formation (Lissauer et al., 2009). Another possible explanation is that metal-richer stars are more likely to form multiple giant planets closer in, and that our sample of single giants contains additional undetected giant companions beyond 30 au. This is disfavored by the results from Fulton, Rosenthal, et al., 2021, which found evidence that giant planets are less common beyond 8 au than within 1–8 au.

In light of our single-multi metallicity result, we explore whether the eccentricities

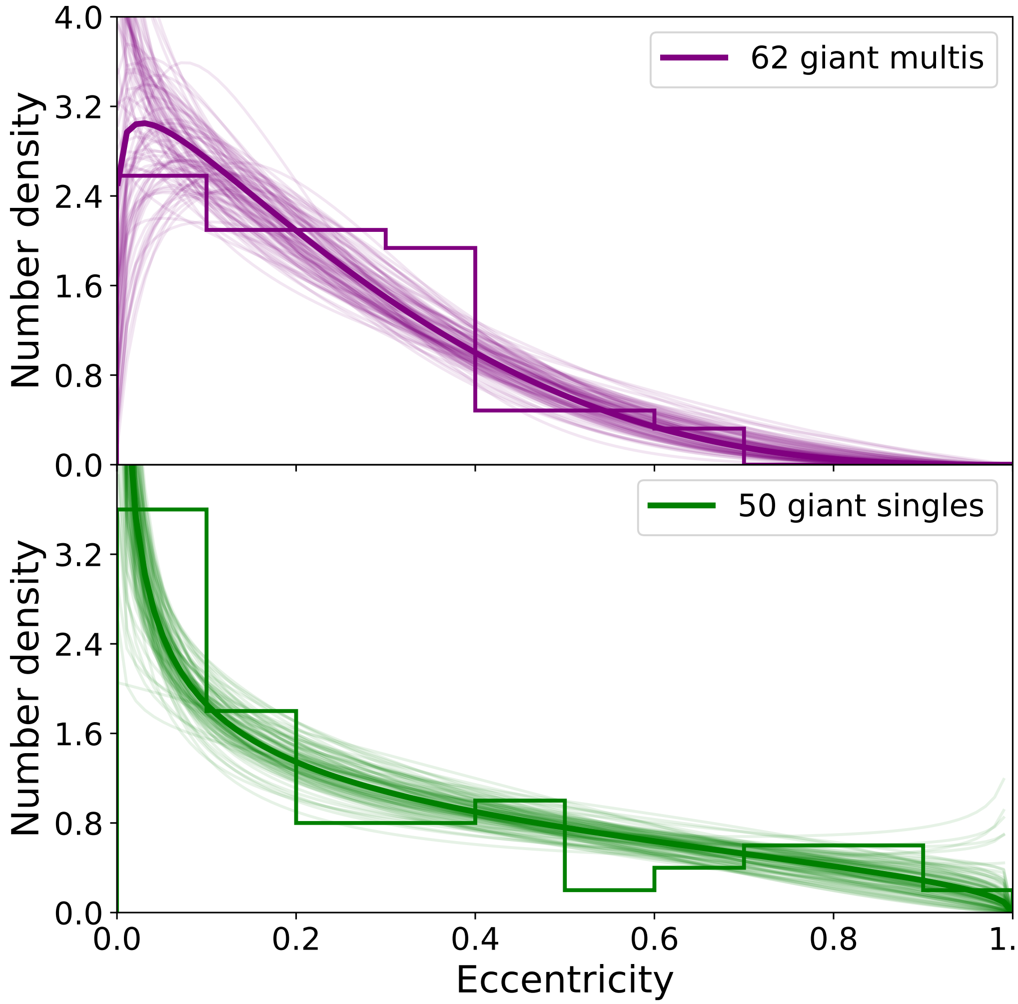

Figure 5.3: Eccentricity distributions of single and multiple giants beyond 0.3 au. Purple/top distribution is for multiple giants, green/bottom distribution is for single giants. Histograms are distributions of individual maximum posterior values, while curves show the median and many draws from beta distribution posteriors, produced using hierarchical inference.

or semi-major axis distribution of the giant planet population is also metallicity-dependent. The hosts of circular ($e < 0.2$) giant planets and hosts of eccentric ($e \geq 0.2$) giant planets have median metallicities separated by less than 1σ in their standard errors, and Kolmogorov-Smirnov tests produce p-values almost entirely above 0.05. The same is true of hot ($a < 0.1$) giant hosts with respect to colder ($a \geq 0.1$) giant hosts. This rules out the possibility that our single-multi metallicity result is driven by a or e as confounding factors, rather than multiplicity itself.

Splitting single giants along $e \geq 0.5$, further along the long tail of eccentric singles,

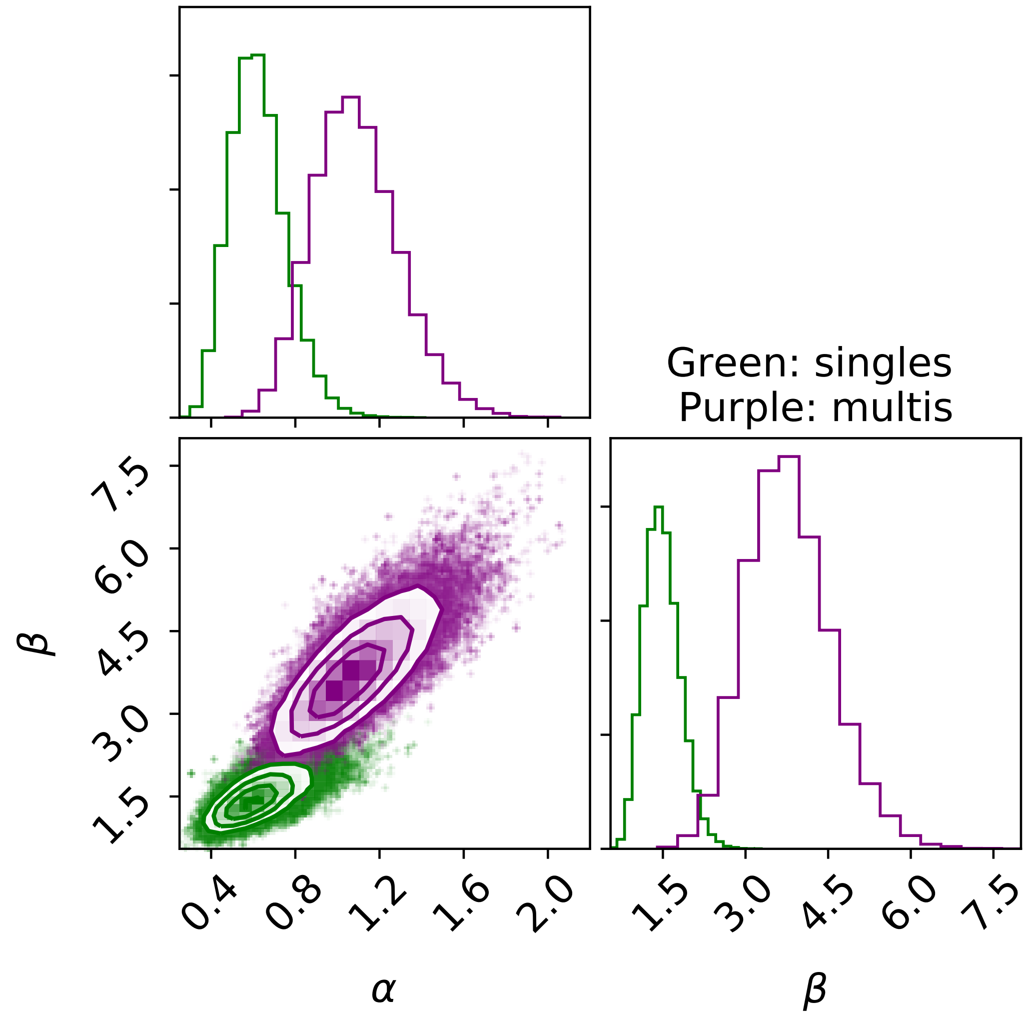

Figure 5.4: Beta distribution posteriors for the single and multiple giant eccentricity distributions. The two populations are $\sim 3\sigma$-distinct.

leads to inconclusive results. Figure 5.7 implies that there is not a conclusive difference in host metallicity between single $e \geq 0.5$ hosts and hosts of less eccentric singles, and Kolmogorov-Smirnov tests also show no conclusive difference. The medians of the two groups are separated by just 1.1σ. However, the same is all true of the difference between single $e \geq 0.5$ hosts and all multi-giant hosts; Kolmogorov-Smirnov tests are inconclusive, and the medians of the two groups are separated by 1σ. We therefore conclude that our sample of single eccentric planets is not large enough to differentiate its metallicity distribution from that of either the population of single planets on circular orbits or the multi-giant population. This is not surprising, as our sample only contains 12 single planets with eccentricities higher than 0.5. The two comparison samples are proportionally larger, with 52 systems with single giants on low eccentricity orbits ($e \leq 0.5$) and 31 systems with multiple giant planets.

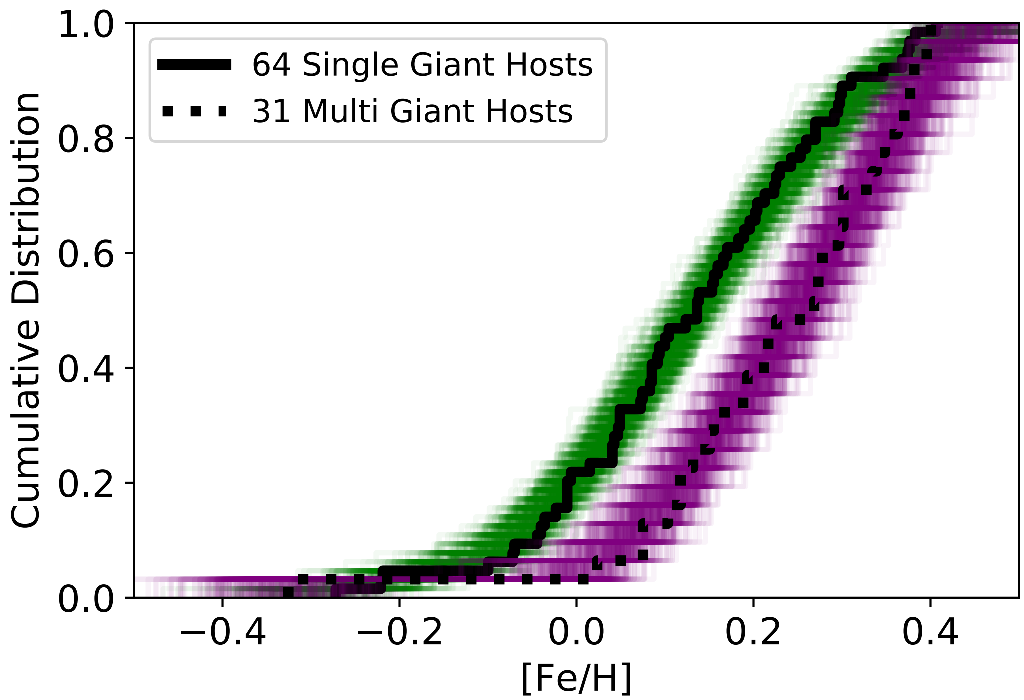

Figure 5.5: Cumulative metallicity distributions for hosts of single giant planets and hosts of multiple giant planets. Solid and dashed steps show distributions of median metallicity measurements, while transparent steps show metallicities drawn many times from Gaussian distributions, with means and standard deviations taken from measurement means and uncertainties. Kolmogorov-Smirnov tests show that the two groups are distinct.

Hot Jupiter pile-up may be missing from multis

Figure 5.8 shows our model for occurrence with respect to orbital separation for the distinct single giant and multiple giant populations. There is a hot Jupiter pile-up within 0.6 au for lonely giants, but no such pile-up for neighborly giants. We further investigate this phenomenon by generating occurrence posteriors for three classes of giants: warm Jupiters (0.1–0.4 au), moderately hot Jupiters (0.06–0.1 au), and very hot Jupiters (0.03–0.06 au). The single giant and multiple giant distributions of warm and hot Jupiters are indistinguishable, but the very hot Jupiter distributions are separated with 97.6% confidence. We measured this confidence as the fraction of random draws from the difference between distributions that is greater than zero.

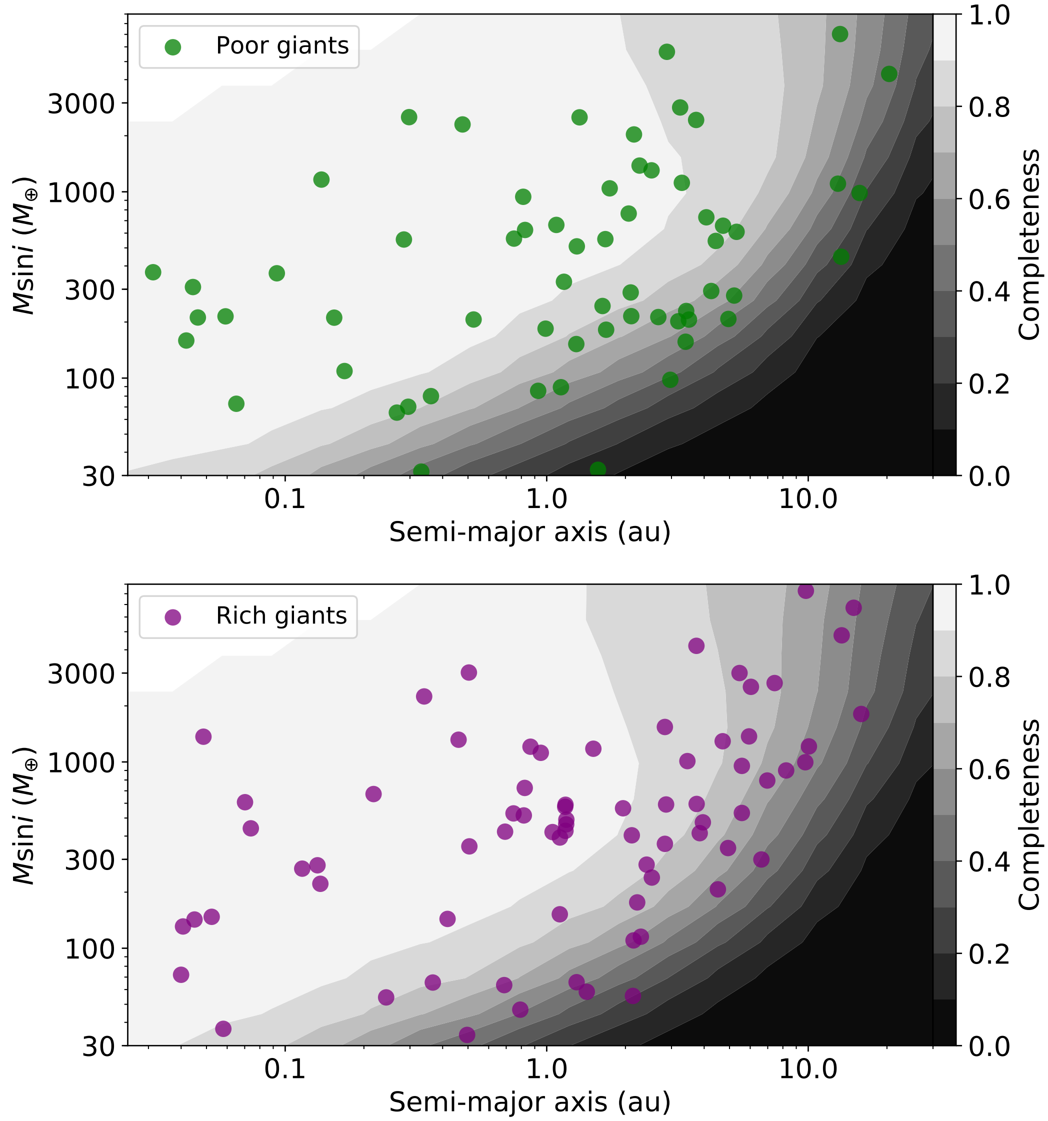

Figure 5.6: Top: Average search completeness contours for giant hosts below their median metallicity. Bottom: Average search completeness contours for giant hosts above their median metallicity. The two contours are nearly equivalent, implying that sensitivity bias to metal-rich stars is not the cause of the single-multi host distinction seen in Figure 5.5.

When we define multiplicity to include all massive companions, including brown dwarfs and stellar binaries, one very hot Jupiter changes in multiplicity status, and our separation confidence decreases to 94.4% significance.

While very hot Jupiters may be lonely, warm Jupiters ($\geq$ 0.1 au) may tell another story. Jackson et al., 2021 predicts that warm Jupiters have outer giant companions with significantly non-zero frequency, which could imply that these companions

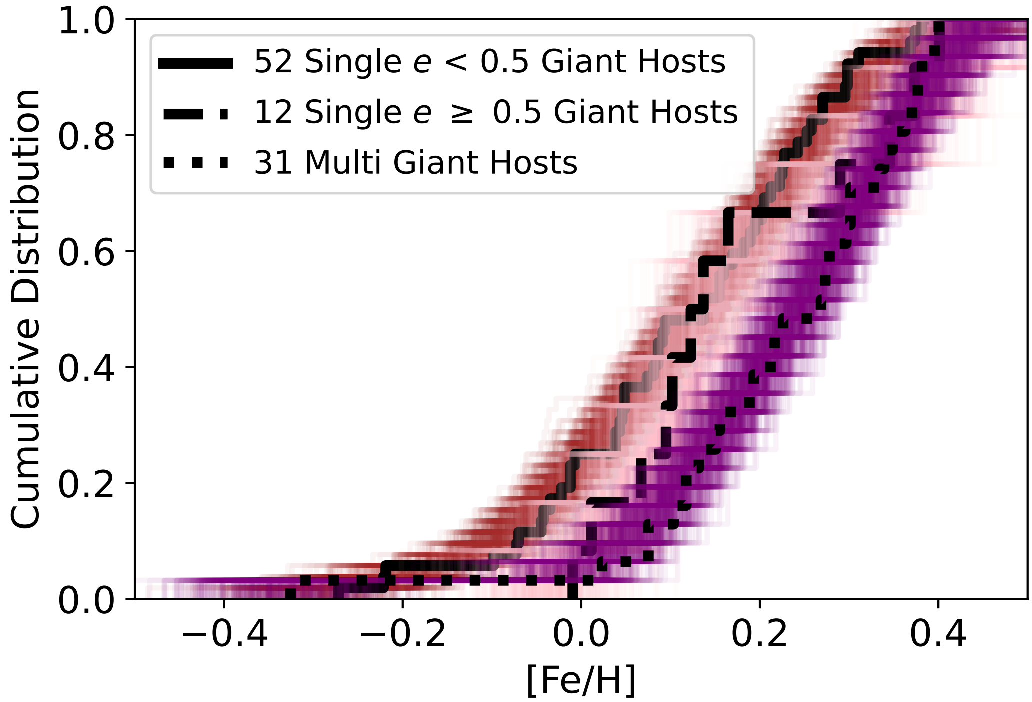

Figure 5.7: Cumulative metallicity distributions for hosts of circular ($e < 0.5$) giant planets, hosts of eccentric ($e \geq 0.5$) giant planets, and multi-giant hosts. See Figure 5.5 for description. The eccentric single hosts are inconclusively drawn from the same underlying metallicity distribution as the less eccentric singles or the multis.

interact secularly to bring the warm Jupiters to their final states. Alternatively, Huang et al., 2016 found that warm Jupiters are more likely to have small planet companions than hot Jupiters. This could imply that warmer giants are less likely to have migrated inward, since migration would disrupt the orbits of small companions. We test these hypotheses with our sample of warm Jupiters, shown in Figure 5.9. In a sample of 21 systems, 10 show detected outer giant companions, while 11 do not. We satisfy the prediction of Jackson et al., 2021, and therefore cannot rule out secular migration. Further resolving this question requires a sample that is uniformly sensitive to both cold gas giants and small rocky planets, which is not the case for the CLS.

Inconclusive mass functions

By investigating occurrence as a function of mass for single and multiple giant planets separately, we may be able to infer whether the two groups form via different pathways. Figure 5.10 shows occurrence rates with respect to $M \sin i$ of both single

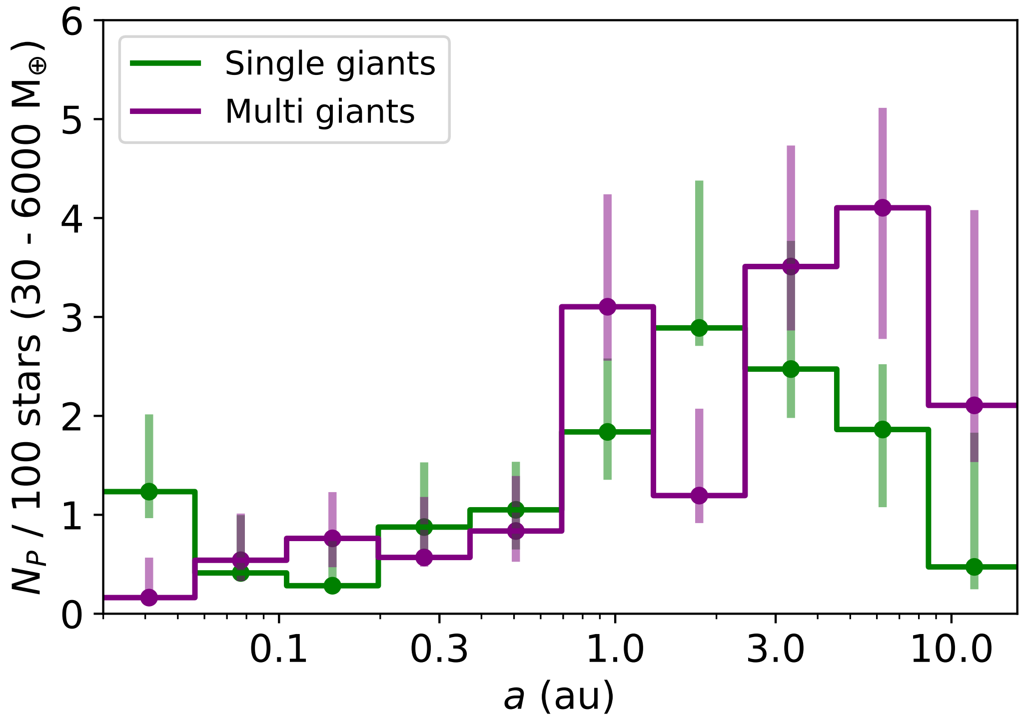

Figure 5.8: Occurrence with respect to semi-major axis for single giant planets and multiple giant planets. Steps and dots are median values, while vertical lines are 15.9–84.1% confidence intervals. The single giant planets display a pile-up of very hot Jupiters within 0.06 au, tentatively not shared by the multiple giant planets.

and multiple giants within 0.1–2 au (multis can include planets outside this range). These two distributions are $\leq 2\sigma$ separated in all four considered ranges of $M \sin i$, but the multi distribution tends toward higher masses than the single distribution. Leaving aside our occurrence model, the median $M \sin i$ of our single giant planets is 0.92 M_J, while the median $M \sin i$ of giant multiples is 1.71 M_J. While we cannot assume that the giant masses are normally distributed, we can report that the standard error in the mean $M \sin i$ for singles is 0.33 M_J, and that the standard error in the mean $M \sin i$ for multis is 0.42 M_J.

Stepping back from our occurrence model, we can compare the cumulative distributions of the two groups and make assumptions that the average completeness contours for the single hosts and multi hosts are nearly identical, similar to the metallicity-split contours in Figure 5.6. Figure 5.11 shows these cumulative distributions. KS tests show that there is a $> 99.7\%$ chance of a p-value below 0.05 for matching single and multi giants within 3 au. This can be interpreted as the degree of confidence in the claim that giants in multiples are consistently more massive

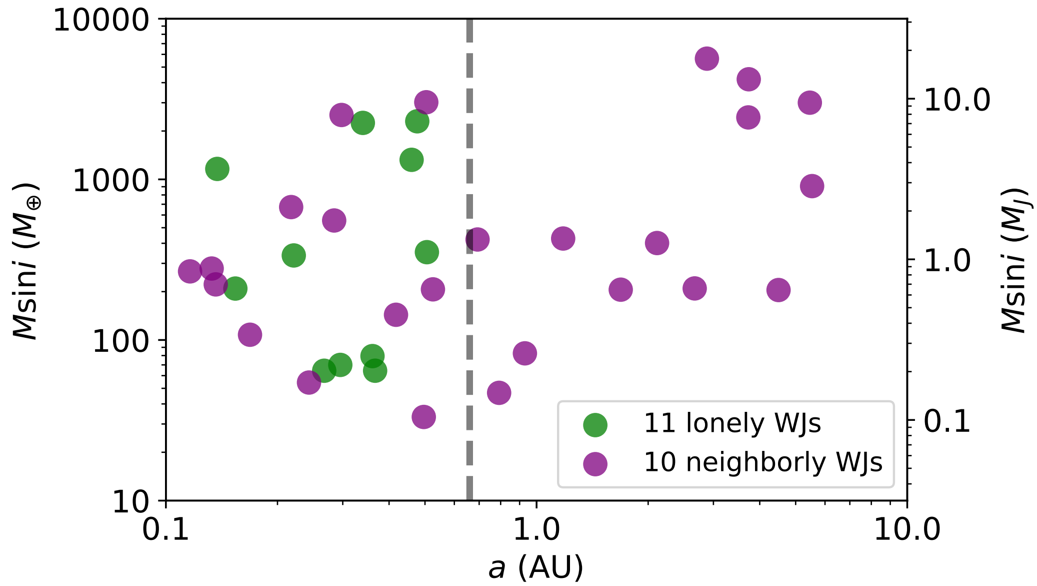

Figure 5.9: Lonely and neighborly warm Jupiters, described in Subsection 5.4. 11 systems have lonely warm Jupiters, while 10 systems host both a warm Jupiter and at least one detected outer giant companion. This does not refute the synthetically driven claim by Jackson et al., 2021 of secular warm Jupiter migration.

than single giants. Additionally, while eccentric singles appear to be more massive than circular singles, this could be a selection effect of more eccentric detections with greater mass.

The properties of multiple giant planets

Comparative work requires taking a broader view of giant multiplicity, beyond hot or warm Jupiters. Figure 5.12 shows the observed giant multiplicity distribution of the CLS. Approximately two thirds of all stars that host a giant planet host only one observed giant planet, while one third of giant hosts show two or more observed giants. This distribution is not corrected for search completeness; it only shows the multiplicity of detected giant planets. This means that at least one third of stars that host one giant planet also host two or more giant planets.

By looking at systems of multiple giant planets, we can also explore how important global disk properties are for setting final planet masses compared to the timeline and location of each individual planet's formation. Focusing on multiple giant systems, we investigate whether paired giant planets have correlated masses. Figure 5.13 compares $M \sin i$ values of neighboring giant planets, limited to one giant pair per system. We measure a Pearson correlation coefficient of $R = 0.45 \pm 0.010$ (error

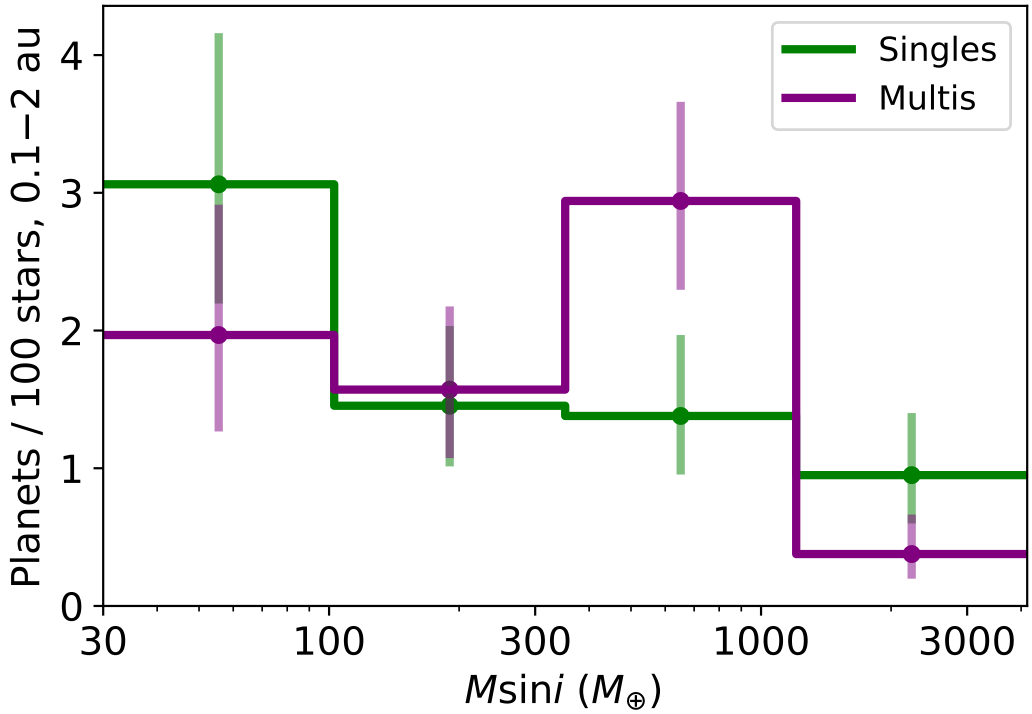

Figure 5.10: Occurrence as a function of $M \sin i$ for single and multiple giants. Steps are median values, while vertical lines are 15.9–84.1% confidence intervals. The two distributions are not $\leq 2\sigma$ separated in all four considered ranges.

come from bootstrapping measured $M \sin i$ values with uncertainties). We use two independent methods to investigate whether sampling a population of uncorrelated giant pairs may produce this value of R. In one case, we use bootstrapping to approximate uncorrelated populations using the observed CLS catalog. First, we repeatedly draw two giant planets from the CLS catalog until we have thirty-one pairs, the same size as the number of observed multi-giant systems. We sample without replacement to ensure that we are shuffling our entire giant catalog, rather than drawing the same planets multiple times. Then, we measure the Pearson correlation coefficient of this shuffled population. We also measure the standard deviation in the log-ratio of planet pair $M \sin i$, and assume coplanarity to simplify this to $\sigma_{\ln(M_2/M_1)}$.

We repeat this process 10^5 times, until we build up bootstrapped distributions of R and $\sigma_{\ln(M_2/M_1)}$. Our measured R falls just within the 99.7% confidence interval of the bootstrapped distribution, at the 99.4% confidence interval. This means that there is a 0.6% probability that an uncorrelated giant planet population could produce a Pearson correlation coefficient of 0.45. Our measured $\sigma_{\ln(M_2/M_1)}$ falls on

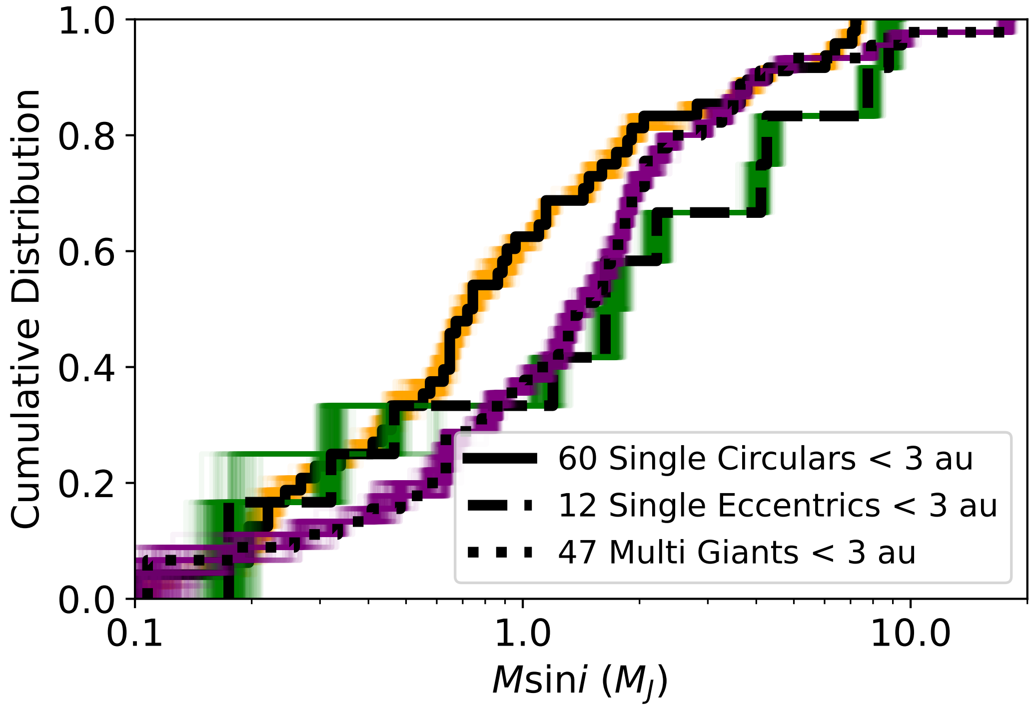

Figure 5.11: Cumulative $M \sin i$ distributions for multiple giants within 3 au and single giants split across an eccentricity of 0.4. While the single and circular giants are $> 4\sigma$ separated from the multis, eccentric giants are inconclusively more massive. Note that this may be a selection effect for more eccentric detections with greater mass.

the 1.4% confidence interval of the bootstrapped distribution.

We also use a synthetic population experiment to independently measure the significance of our observed correlation. We drew giant planet pairs from a uniform $\ln(M \sin i\ )$ distribution and the semi-major axis distribution reported in Fulton, Rosenthal, et al., 2021, with a minimum period semi-major axis ratio $\ln(a_2/a_1) >$ 0.5. We impose this limit on the semi-major axis ratio in order to prevent unlikely or unphysical draws of pairs within 3:2 resonance. We took our search completeness contours into account to only select simulated planet pairs for which both planets are detected. We detail our synthetic population algorithm in the list below.

- Draw two values of $\ln(M \sin i)$ from a uniform distribution between $\ln(30\ M_\oplus)$–$\ln(6000\ M_\oplus)$, and two values of $\ln(a)$ from the nonparametric giant planet model reported in Fulton, Rosenthal, et al., 2021, bounded between 0.1–10 au. We approximate this distribution as piecewise-uniform in $\ln(a)$,

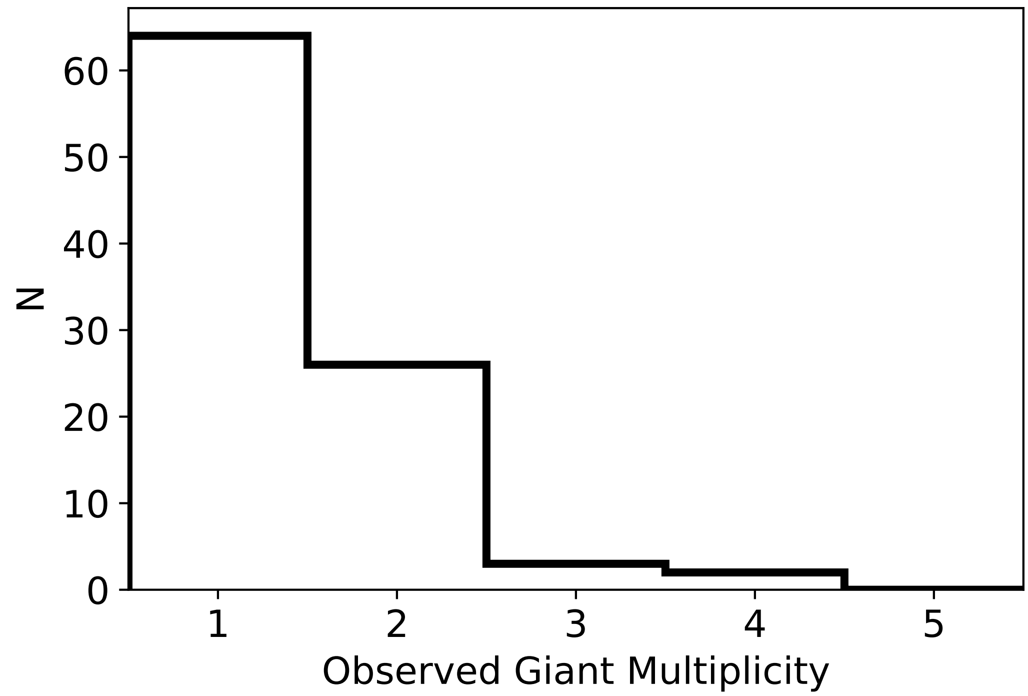

Figure 5.12: Observed multiplicity of giant planets in the CLS sample. Half as many systems contain two or more giants as contain only one giant.

with a break at 1 au, integrated probability of 2/11 between 0.1–1 au, and integrated probability of 9/11 between 1–10 au.

- Measure the search completeness and draw a random number from a uniform 0–1 distribution for each set of $\ln(a)$ and $\ln(M \sin i)$. For each set, if the random number is less than the search completeness, report the planet corresponding to that set as detected.
- Keep the planet pair if we detect both planets and $\ln(a_2/a_1) > 0.5$, otherwise reject it.
- Repeat until we have drawn the desired number of simulated planet pairs for one synthetic population.
- Repeat until we have drawn the desired number of synthetic populations.

With 10^4 synthetic populations of 31 observed planet pairs each, this test finds that our observed correlation falls just within the 99% confidence interval of the distribution that our uncorrelated simulations produces.

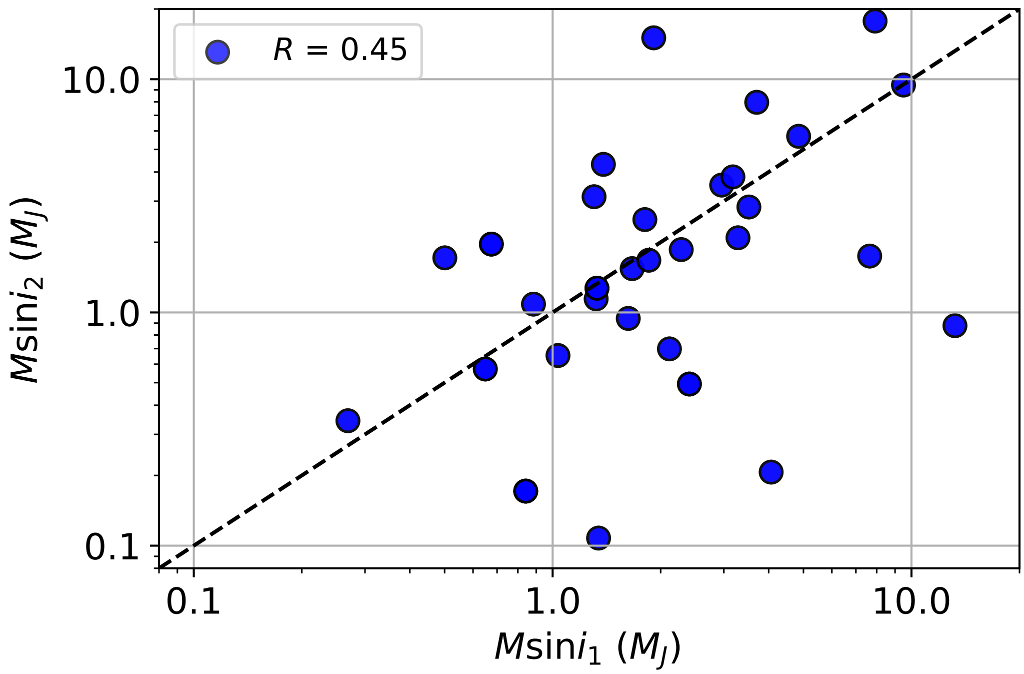

Figure 5.13: $M \sin i$ distributions of giant planet pairs. We observe a 0.45 correlation in $\ln(M \sin i)$ between neighboring giants.

5.5 Discussion

Metallicity, eccentricity, and multiplicity

Dawson and Murray-Clay, 2013 found that giants with metal-rich hosts show signatures of planet-planet scattering. In particular, giant planets between 0.1 AU and 1 AU are likely to be more eccentric when they orbit stars with [Fe/H] > 0. This higher eccentricity could be caused by planet-planet scattering events, which would imply that metal-rich stars are more likely to host multiple giant planets. Our results support this hypothesis, as we find that systems of multiple giant planets are more common around metal-rich stars, and that systems with multiple giant planets have higher average eccentricities than those in single giant planet systems. Table 5.1 shows the above properties side-by-side for the two populations.

Property	Trait	Single Giant Planets	Multiple Giant Planets
Eccentricity			
	$e_{10\%}$	0.03	0.03
	$e_{50\%}$	0.20	0.23
	$e_{90\%}$	0.77	0.47

	Beta α	$0.60^{+0.14}_{-0.12}$	$1.06^{+0.21}_{-0.18}$
	Beta β	$1.45^{+0.34}_{-0.30}$	$3.68^{+0.81}_{-0.70}$
Semi-major Axis			
	Feature	HJ pile-up	HJ deficit
Mass			
	Median	0.92 M_{J}	1.71 M_{J}
	Mean	2.07 ± 0.33 M_{J}	2.88 ± 0.42 M_{J}
	Feature	Less massive	Pair uniformity
Iron Abundance			
	< [Fe/H] >	0.129 ± 0.019	0.228 ± 0.027
	Feature	Enriched over Solar	More enriched
Stellar Binarity			
	Fraction	$26.2 \pm 6.3\%$	$25.8 \pm 9.1\%$

Table 5.1: Summary of population comparison between single giant systems and multi giant systems. Fraction of bound binaries comes from a literature estimate with Simbad and the Washington Double-star Catalog. Uncertainties come from Poisson counting error.

Although this basic story seems clear, our data are less definitive on the question of how the sub-population of single giants on highly eccentric orbits acquired their eccentricities. There are few other mechanisms that could explain the metal-poor, high-eccentricity tail of lonely giants above $e \sim 0.6$. The most likely alternative to scattering is the Kozai-Lidov mechanism, by which a massive outer companion secularly perturbs the orbit of an inner companion, can potentially excite giants to $e > 0.6$, but only given a high mutual inclination and absence of other dynamical factors Naoz, 2016. Even if this could account for all of our single, highly eccentric giants, it would require the presence of stellar binary companions or undetected giant companions in these systems. We went back and examined the radial velocity data sets for the eccentric single planets in our sample, and found that 4 out of 65 single giant systems have $\geq 3\sigma$-significant parabolic or linear trends, and two of these (HD 34445 and HD 156668) show strong linear correlation between their RV trends and S-values, which we can treat as a proxy for magnetic activity. Therefore, only 2 out of 65 single giant systems show evidence for a stellar binary companion (HD 195019 and HD 145934), and neither of these systems host giants with $e \geq 0.2$. HD 195019 has a visually resolved wide-orbit binary companion (Lu et al., 1987), while HD 145934's status is unclear.

Alternatively, it is possible that most of the single planets in our sample do in fact have additional sub-Jovian giant companions located outside 10 au. Our survey may be missing giant planets beyond this distance, and therefore misidentifying multi-giant systems as single giant systems. As seen in Figure 5.14, the majority of CLS-detected giant planet pairs have a semi-major axis ratio less than 10, with a peak ratio frequency around 3. This means that our coldest lonely giants around 10–20 au could have companions out around 30–60 au, where they would be undetectable if they were small enough and on circular orbits. Again, this is disfavored by the giant planet fall-off beyond 8 au reported in Fulton, Rosenthal, et al., 2021.

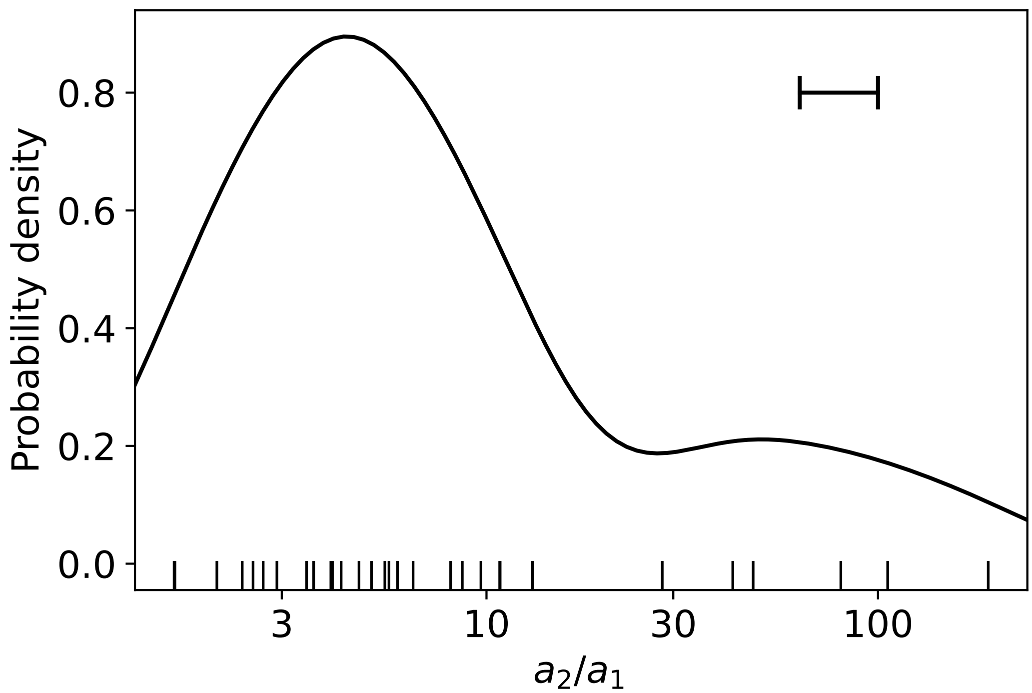

Figure 5.14: Observed semi-major axis ratio distribution of two-giant systems. The curve is a kernel density estimation of the distribution, while the vertical ticks show the observed set of axis ratios. Horizontal bar shows the kernel width.

Where do our lonely and highly eccentric giants come from? Out of 50 lonely CLS giant planets beyond 0.3 au, 11 have median eccentricities greater than 0.5. CLS contains 30 multi-giant systems with at least one giant beyond 0.3 au. If the eccentric singles all used to be scattered multis, that would imply that 11 out of 41 multi systems, or roughly a quarter, scattered themselves into singledom. Carrera et al., 2019 found that the observed eccentricity distribution of giants within 5 au is compatible with scattering accounting for all giants with $e \geq 0.3$. This supports the claim that many highly eccentric singles used to be multis that experienced scattering events. However, our poorly constrained metallicity distribution for high-eccentricity giant hosts limits our ability to make claims about their original multiplicity.

We leave this discussion with more questions than we entered with, but are hopeful that future work will elucidate these mysteries. Longer RV baselines for eccentric single systems and imaging campaigns to look for distant stellar or substellar companions could clarify the statuses of our lonely and eccentric giants, while synthetic population models might have something interesting to say about the origins of our observed eccentricity and metallicity distributions.

Ultra-hot Jupiter pile-up

Knutson et al., 2014 performed an RV-trend search for massive companions to hot Jupiters, and found that $50 \pm 10\%$ of giants within 0.7–11 days host an outer companion between 1–13 M_{J} and 1–20 au. More than two-thirds of their hot Jupiters with companions have orbital periods within 6 days. In the CLS sample, 3 out of 15 hot Jupiters within 10 days host an outer companion in the same range (HD 187123, HD 217107, HD 9826). That gives a lower bound on the HD companion rate of $20 \pm 12\%$. These two rates are compatible within 2.14σ, different but not definitively so.

Still, why do we see fewer companions to ultra-hot Jupiters in the CLS than in this other study? One possible explanation is contamination of RV data by stellar magnetic activity. Three of the systems in Knutson et al., 2014, specifically HAT-P-4, HAT-P-22, and HAT-P-32, show strong correlation with the core strength of the Calcium H & K line, which is a reliable indicator of stellar activity. The observational baselines of their datasets with detected trends range from 5 to 8 years, so 11 out of 14 reported outer companions are modeled as linear trends. This makes it more challenging to differentiate between stellar activity and planetary signals. We have obtained additional observations of these stars over the last four years, doubling some of the baselines and strengthening the activity correlations for these three stars (conversations with Prof. Heather Knutson). Beyond activity, another RV measurement of XO-2 taken in 2021 indicates that there is no significant trend in that RV dataset. Taken together, these potential false positives would reduce the estimated hot Jupiter companion rate from this study by roughly a third.

If this single hot Jupiter enhancement over multiples is indeed real, it could support the claim that most many Jupiters form *in situ*, rather than forming at large distances from their host stars and then migrating inward (Batygin et al., 2016). The presence of an outer giant could suppress the transfer of material from further out in the disk onto a hot protoplanet, which would no longer have the gas supply needed to become a giant.

Giant peas in a pod

Previous studies of sub-Neptune-sized transiting planets have found that planets located in the same system have correlated masses and radii (Weiss et al., 2018; Millholland, Wang, et al., 2017). Our study reveals that giant planets forming in the same system also appear to have correlated masses. This could imply that the global properties of protoplanetary disks strongly influence the final masses of their nascent giant planets, and that these properties are more important than solid surface density and other properties that vary radially or over time. That condition could be initial protoplanetary disk mass, since disk mass determines how much material is available during giant planet formation, solid mass for large core formation, or disk lifetime, since quickly dissipating disks have less time to make giants.

Weiss et al., 2018 measured a Pearson correlation of $R = 0.65$ for the paired planet radii of a larger sample than the CLS. This means that the magnitude of the 'small peas in a pod' effect is significantly greater than that of the 'giant peas in a pod' effect. The final masses of giant planets are very sensitive to the relative timing of the onset of runaway accretion. Giant planets can also cut of the flow of gas to the inner disk, which could affect the ability of an inner companion to accrete a massive gas envelope. This might explain the weaker correlation between giant masses.

5.6 Conclusions

We compared and contrasted single giant systems and multi-giant systems, and found several key differences. When we look at giant occurrence with respect to orbital separation, giant singles have a super-hot Jupiter pile-up not shared by giant multis, limited by lack of sensitivity to giant planets beyond 20 au. We reproduced previously seen eccentricity distributions for these two groups, and found that the lonely giants have a circular orbit pile-up and long eccentric tail, while the multis are spread across moderate eccentricities and limited to $e \leq 0.7$. We also found that multi-giant stellar hosts are more metal-rich than single giant hosts, and that total giant planet mass (inferred from $M \sin i$) does not correlate with host metallicity. Taken together, these results may imply that more metal-rich stars form more discrete giant planets, possibly due to formation of more rocky cores. We also find that at least a third of stars that host giant planets host two or more of them. This implies that giant planet pairs probably have correlated masses, and perhaps that the shared formation environment of giants influences their final mass limits.

Although this work provides new information about the underlying population of giants, it also raises unresolved questions about why we see these patterns. What does the lonely super-hot Jupiter pile-up mean for theories of hot Jupiter formation? Are lonely and highly eccentric giants products of planet-planet scattering? And why are giant planet pair masses correlated? These questions can be the starting points for new theoretical inquiries into planet formation.

All code used in this chapter is available at `github.com/California-Planet-Search/rvsearch` and `github.com/leerosenthalj/CLSIV`. This research makes use of GNU Parallel (Tange, 2011). We made use of the following publicly available Python modules: `astropy` (Astropy Collaboration et al., 2013), `matplotlib` (Hunter, 2007), `numpy/scipy` (Walt et al., 2011), `pandas` (McKinney, 2010), `emcee` (Foreman-Mackey, Hogg, Lang, et al., 2013a), and `RadVel` (Fulton, Petigura, Blunt, et al., 2018).

Chapter 6

CONCLUSIONS

In this thesis, I have investigated the population statistics and formation mechanisms of a myriad of planet types. Chapter 2 presents a new planet catalog ranging from hot super-Earths to cold Jovians, and the following chapters explore everything in-between. Although studies of cold Jovians, rocky planets, and multi-planet systems may seem unrelated, they share a unifying theme: the Solar System. Each chapter of this thesis touches a different facet of our cosmic home and tells us something about how we do or do not fit in among our exoplanetary neighbors. Chapter 3 reports on the occurrence of gas giants as a function of orbital distance, and shows that Jupiter and Saturn are both found in high-occurrence orbits. In other words, the Solar System has typically placed giant planets.

Chapter 4 considers the relationship between close-in super-Earths and outer gas giants, and while the Earth is less massive and further out than any small planets in the CLS sample, this work can still tell us something about the Solar System if we extrapolate from super-Earths to Earth analogs. Assuming that the two populations have similar relationships to giant planets, we can again infer that the Solar System is typical, this time with respect to having giant planets that are cold enough to not suppress small planet formation. This is a speculative conclusion, but it is physically intuitive that if Jupiter were in a 2 au orbit instead of a 5 au orbit, the Earth would be less likely to have a stable formation timeline.

Chapter 5's exploration of the difference between single giant and multiple giant systems shows that these two groups have different metallicity distributions. This reveals one way in which the Solar System is *atypical*: the Sun is far more metal-poor than most stars that host several giants. Fischer and Valenti, 2005 already established that the Sun is anomalously poor as a giant planet host, and now it seems that it is even more anomalous as a host of both Jupiter and Saturn. This may be a statistical fluke, but it prompts curiosity about the Sun and how it might be an oddity among its planet-hosting peers.

What is the best way to use these results and the rest of this thesis to take new steps in exoplanet research? One path involves creating theoretical models to explain some of the novel findings reported here. There are not yet clear explanations for the 'giant peas-in-a-pod' effect, for giant planets that both are highly eccentric and appear to be lonely, or for our observed small planet conditional and absolute occurrence rates. Solving these puzzles with simulations or analytical methods could add details to the unfinished story of planet formation.

Equally or even more importantly, these results must be independently reproduced in order to be taken more seriously. There have been a number of replication crises across academia in the past few decades, and astronomy is not immune to this issue. Since it is possible that some of my data, methods, or reasoning are flawed, I would be delighted to see other RV surveys affirm or challenge any of the above work. There is a substantial amount of archival HARPS RV data available for stars in the southern hemisphere (Trifonov et al., 2020) with little CLS overlap. With some careful sample selection, this dataset could be an excellent starting point for exploring the above questions with a different survey.

Of course, the Holy Grail of exoplanet research for the foreseeable future is the detection and characterization of Earth twins. In order to make this goal feasible without acquiring several thousands of observations per star, state-of-the-art RV precision needs to advance from ~60 cm s^{-1} as of the publication of this thesis to ~10 cm s^{-1}. This requires progress on several fronts, as total RV precision is the sum of noise contributions by instrument noise and systematics, stellar activity, and imperfections in Doppler reduction pipelines. First, instruments must become sufficiently precise, wavelength-calibrated, and temperature-stable that thermal and instrumental noise do not dominate over Doppler shifts caused by Earth twins. Technological developments such as laser frequency combs, temperature-stable materials, and improved optics and software are making their way into the next generation of RV instruments. HARPS-3 (Thompson et al., 2016) and the Keck Planet Finder (KPF; Gibson et al., 2018) are just some of the spectrographs under development with predicted instrumental sensitivities near 30 cm s^{-1}, coming closer to detecting an Earth twin orbiting a G dwarf.

Instrument improvements are necessary but insufficient for the discovery of Earth twins. The other key obstacle is the influence of stellar activity on spectra. Both years-long fluctuations in magnetic activity and spot rotation on timescales of weeks to months change the shapes of stellar absorption lines and can cause both quasi-periodic and white noise in Doppler measurements (Dumusque et al., 2011; Robertson, Mahadevan, et al., 2014). Such noise both obscures weak RV signals and creates false-positive planet detections. There are two clear paths to circumventing this issue. One can deal with existing noise in Doppler measurements by using correlated noise models, such as Gaussian processes (Rajpaul et al., 2015). Alternatively, one can mitigate the effect of activity on extracting Doppler measurements from spectra. This is possible through accurate models of line profiles, or by masking out lines that are more strongly affected by activity (Dumusque, 2018). All of the above require higher-cadence data to capture all relevant timescales of activity, higher signal-to-noise measurements, and instrumental improvements in spectral resolution (Zhao et al., 2022).

Besides advances in instrumentation and modeling, carefully targeted and conducted surveys are a key component in searching for other Earths. The CLS is an achievement that required decades of work, but it is not the product of systematic and consistent observing. Cadences for individual stars varied widely over the years, resulting in heterogeneous sensitivity to planet detections from star to star. The one planned survey with consistent and high-cadence observing is the Terra Hunting Experiment (Hall et al., 2018), which will use HARPS-3 to perform a high-cadence and unbiased search for Earth twins orbiting Solar twins. Other collaborations can and should undertake more targeted searches like this to maximize the probability of discovery. This brings us to the part that this thesis can play in the search for habitable worlds. The CLS is an ideal sample for extreme-precision radial velocity surveys because of the brightness, proximity, and low activity levels of its stars. Additionally, Chapter 4 of this thesis shows that cold giant planets do not suppress inner small planet formation, at least within certain mass ranges. Based on this finding, KPF and future extremely precise RV instruments could conduct targeted searches for Earth twins orbiting known cold giant hosts in the CLS, in an attempt to find true Solar System analogs. The combination of decades of HIRES/Hamilton/APF observations and high-cadence, high-precision data could provide unique sensitivity to rocky planets close to 1 au. Through this or other search programs, and with enough time and patience, humanity might find at least one lookalike neighbor in the Milky Way within the next few decades.

BIBLIOGRAPHY

Alonso, R. et al. (Oct. 2004). "TrES-1: The Transiting Planet of a Bright K0 V Star". In: *ApJL* 613.2, pp. L153–L156. DOI: 10.1086/425256. arXiv: astro-ph/0408421 [astro-ph].

Apps, K. et al. (Feb. 2010). "M2K: I. A Jupiter-Mass Planet Orbiting the M3V Star HIP 79431". In: *PASP* 122.888, p. 156. DOI: 10.1086/651058. arXiv: 1001.1174 [astro-ph.EP].

Astropy Collaboration et al. (Oct. 2013). "Astropy: A community Python package for astronomy". In: *A&A* 558, A33, A33. DOI: 10.1051/0004-6361/201322068. arXiv: 1307.6212 [astro-ph.IM].

Bakos, G. Á. et al. (Sept. 2002). "System Description and First Light Curves of the Hungarian Automated Telescope, an Autonomous Observatory for Variability Search". In: *PASP* 114.799, pp. 974–987. DOI: 10.1086/342382. arXiv: astro-ph/0206001 [astro-ph].

Baranne, A. et al. (Oct. 1996). "ELODIE: A spectrograph for accurate radial velocity measurements." In: *A&AS* 119, pp. 373–390.

Batygin, K. (Aug. 2015). "Capture of planets into mean-motion resonances and the origins of extrasolar orbital architectures". In: *MNRAS* 451.3, pp. 2589–2609. DOI: 10.1093/mnras/stv1063. arXiv: 1505.01778 [astro-ph.EP].

Batygin, K., P. H. Bodenheimer, and G. P. Laughlin (Oct. 2016). "In Situ Formation and Dynamical Evolution of Hot Jupiter Systems". In: *ApJ* 829.2, 114, p. 114. DOI: 10.3847/0004-637X/829/2/114. arXiv: 1511.09157 [astro-ph.EP].

Benatti, S. et al. (July 2020). "The GAPS programme at TNG. XXIII. HD 164922 d: close-in super-Earth discovered with HARPS-N in a system with a long-period Saturn mass companion". In: *A&A* 639, A50, A50. DOI: 10.1051/0004-6361/202037939. arXiv: 2005.03368 [astro-ph.EP].

Biller, B. A. et al. (Nov. 2013). "The Gemini/NICI Planet-Finding Campaign: The Frequency of Planets around Young Moving Group Stars". In: *ApJ* 777.2, 160, p. 160. DOI: 10.1088/0004-637X/777/2/160. arXiv: 1309.1462 [astro-ph.EP].

Bitsch, B., S. N. Raymond, and A. Izidoro (Apr. 2019). "Rocky super-Earths or waterworlds: the interplay of planet migration, pebble accretion, and disc evolution". In: *A&A* 624, A109, A109. DOI: 10.1051/0004-6361/201935007. arXiv: 1903.02488 [astro-ph.EP].

Bitsch, B. and S. Savvidou (Jan. 2021). "Influence of grain sizes and composition on the contraction rates of planetary envelopes and on planetary migration". In: *arXiv e-prints*, arXiv:2101.03818, arXiv:2101.03818. arXiv: 2101.03818 [astro-ph.EP].

Bitsch, B., T. Trifonov, and A. Izidoro (Nov. 2020). "The eccentricity distribution of giant planets and their relation to super-Earths in the pebble accretion scenario". In: *A&A* 643, A66, A66. DOI: 10.1051/0004-6361/202038856. arXiv: 2009.11725 [astro-ph.EP].

Blunt, S. et al. (Nov. 2019). "Radial Velocity Discovery of an Eccentric Jovian World Orbiting at 18 au". In: *AJ* 158.5, 181, p. 181. DOI: 10.3847/1538-3881/ab3e63. arXiv: 1908.09925 [astro-ph.EP].

Bonfils, X. et al. (Jan. 2013). "The HARPS search for southern extra-solar planets. XXXI. The M-dwarf sample". In: *A&A* 549, A109, A109. DOI: 10.1051/0004-6361/201014704. arXiv: 1111.5019 [astro-ph.EP].

Borucki, W. J. et al. (Feb. 2010). "Kepler Planet-Detection Mission: Introduction and First Results". In: *Science* 327, p. 977. DOI: 10.1126/science.1185402.

Borucki, W. J. (Mar. 2016). "KEPLER Mission: development and overview". In: *Reports on Progress in Physics* 79.3, 036901, p. 036901. DOI: 10.1088/0034-4885/79/3/036901.

Boss, A. P. (Jan. 1997). "Giant planet formation by gravitational instability." In: *Science* 276, pp. 1836–1839. DOI: 10.1126/science.276.5320.1836.

Bouchy, F. et al. (Mar. 2009). "The HARPS search for southern extra-solar planets. XVII. Super-Earth and Neptune-mass planets in multiple planet systems HD 47 186 and HD 181 433". In: *A&A* 496.2, pp. 527–531. DOI: 10.1051/0004-6361:200810669. arXiv: 0812.1608 [astro-ph].

Bowler, B. P., S. C. Blunt, and E. L. Nielsen (Feb. 2020). "Population-level Eccentricity Distributions of Imaged Exoplanets and Brown Dwarf Companions: Dynamical Evidence for Distinct Formation Channels". In: *AJ* 159.2, 63, p. 63. DOI: 10.3847/1538-3881/ab5b11. arXiv: 1911.10569 [astro-ph.EP].

Bowler, B. P., M. C. Liu, et al. (Jan. 2015). "Planets around Low-mass Stars (PALMS). IV. The Outer Architecture of M Dwarf Planetary Systems". In: *ApJS* 216.1, 7, p. 7. DOI: 10.1088/0067-0049/216/1/7. arXiv: 1411.3722 [astro-ph.EP].

Bryan, M. L., H. A. Knutson, A. W. Howard, et al. (Apr. 2016). "Statistics of Long Period Gas Giant Planets in Known Planetary Systems". In: *ApJ* 821.2, 89, p. 89. DOI: 10.3847/0004-637X/821/2/89. arXiv: 1601.07595 [astro-ph.EP].

Bryan, M. L., H. A. Knutson, E. J. Lee, et al. (Feb. 2019). "An Excess of Jupiter Analogs in Super-Earth Systems". In: *AJ* 157.2, 52, p. 52. DOI: 10.3847/1538-3881/aaf57f. arXiv: 1806.08799 [astro-ph.EP].

Bryson, S. et al. (Jan. 2021). "The Occurrence of Rocky Habitable-zone Planets around Solar-like Stars from Kepler Data". In: *AJ* 161.1, 36, p. 36. DOI: 10.3847/1538-3881/abc418. arXiv: 2010.14812 [astro-ph.EP].

Buchhave, L. A. et al. (May 2014). "Three regimes of extrasolar planet radius inferred from host star metallicities". In: *Nature* 509, pp. 593–595. DOI: 10.1038/nature13254. arXiv: 1405.7695 [astro-ph.EP].

Buchner, J. (June 2016). *PyMultiNest: Python interface for MultiNest*. ascl: 1606.005.

Butler, R. P., G. W. Marcy, E. Williams, et al. (June 1996). "Attaining Doppler Precision of 3 M s-1". In: *PASP* 108, p. 500. DOI: 10.1086/133755.

Butler, R. P., J. T. Wright, et al. (July 2006). "Catalog of Nearby Exoplanets". In: *ApJ* 646.1, pp. 505–522. DOI: 10.1086/504701. arXiv: astro-ph/0607493 [astro-ph].

Butler, R. P., G. W. Marcy, S. S. Vogt, et al. (Jan. 2003). "Seven New Keck Planets Orbiting G and K Dwarfs". In: *ApJ* 582.1, pp. 455–466. DOI: 10.1086/344570.

Butler, R. P., S. S. Vogt, et al. (May 2017). "The LCES HIRES/Keck Precision Radial Velocity Exoplanet Survey". In: *AJ* 153.5, 208, p. 208. DOI: 10.3847/1538-3881/aa66ca. arXiv: 1702.03571 [astro-ph.EP].

Carrera, D., S. N. Raymond, and M. B. Davies (Sept. 2019). "Planet-planet scattering as the source of the highest eccentricity exoplanets". In: *A&A* 629, L7, p. L7. DOI: 10.1051/0004-6361/201935744. arXiv: 1903.02564 [astro-ph.EP].

Cassan, A. et al. (Jan. 2012). "One or more bound planets per Milky Way star from microlensing observations". In: *Nature* 481.7380, pp. 167–169. DOI: 10.1038/nature10684. arXiv: 1202.0903 [astro-ph.EP].

Chachan, Y., P. A. Dalba, et al. (Feb. 2022). "Kepler-167e as a Probe of the Formation Histories of Cold Giants with Inner Super-Earths". In: *ApJ* 926.1, 62, p. 62. DOI: 10.3847/1538-4357/ac3ed6. arXiv: 2112.00747 [astro-ph.EP].

Chachan, Y., E. J. Lee, and H. A. Knutson (Jan. 2021). "Radial Gradients in Dust Opacity Lead to Preferred Region for Giant Planet Formation". In: *arXiv e-prints*, arXiv:2101.10333, arXiv:2101.10333. arXiv: 2101.10333 [astro-ph.EP].

Charbonneau, D. et al. (Jan. 2000). "Detection of Planetary Transits Across a Sun-like Star". In: *ApJL* 529, pp. L45–L48.

Chatterjee, S. et al. (Oct. 2008). "Dynamical Outcomes of Planet-Planet Scattering". In: *ApJ* 686.1, pp. 580–602. DOI: 10.1086/590227. arXiv: astro-ph/0703166 [astro-ph].

Chen, J. and D. Kipping (Nov. 2016). *Chenjj2/Forecaster: First Release Of Forecaster*. Version v1.0.0. DOI: 10.5281/zenodo.164450.

Clanton, C. and B. S. Gaudi (Mar. 2016). "Synthesizing Exoplanet Demographics: A Single Population of Long-period Planetary Companions to M Dwarfs Consistent with Microlensing, Radial Velocity, and Direct Imaging Surveys". In: *ApJ* 819.2, 125, p. 125. DOI: 10.3847/0004-637X/819/2/125. arXiv: 1508.04434 [astro-ph.EP].

Crowe, M. (1986). *The Extraterrestrial Life Debate, 1750–1900: the Idea of a Plurality of Worlds from Kant to Lowell*. Cambridge University Press. URL: https://www.cambridge.org/core/journals/british-journal-for-the-history-of-science/article/abs/michael-j-crowe-the-extraterrestrial-life-debate-17501900-the-idea-of-a-plurality-of-worlds-from-kant-to-lowell-cambridge-cambridge-university-press-1986-pp-xix-680-isbn-0521263050-4000/D176B397E271E700C56D5AA22F3499F6.

Cumming, A. et al. (May 2008). "The Keck Planet Search: Detectability and the Minimum Mass and Orbital Period Distribution of Extrasolar Planets". In: *PASP* 120.867, p. 531. DOI: 10.1086/588487. arXiv: 0803.3357 [astro-ph].

Dalba, P. A. et al. (Mar. 2021). "Giant Outer Transiting Exoplanet Mass (GOT 'EM) Survey. I. Confirmation of an Eccentric, Cool Jupiter with an Interior Earth-sized Planet Orbiting Kepler-1514". In: *AJ* 161.3, 103, p. 103. DOI: 10.3847/1538-3881/abd408. arXiv: 2012.04676 [astro-ph.EP].

Dawson, R. I. and J. A. Johnson (Sept. 2018). "Origins of Hot Jupiters". In: *ARA&A* 56, pp. 175–221. DOI: 10.1146/annurev-astro-081817-051853. arXiv: 1801.06117 [astro-ph.EP].

Dawson, R. I. and R. A. Murray-Clay (Apr. 2013). "Giant Planets Orbiting Metal-rich Stars Show Signatures of Planet-Planet Interactions". In: *ApJL* 767.2, L24, p. L24. DOI: 10.1088/2041-8205/767/2/L24. arXiv: 1302.6244 [astro-ph.EP].

Díaz, R. F. et al. (May 2019). "The SOPHIE search for northern extrasolar planets. XIV. A temperate (T_{eq} 300 K) super-earth around the nearby star Gliese 411". In: *A&A* 625, A17, A17. DOI: 10.1051/0004-6361/201935019. arXiv: 1902.06004 [astro-ph.EP].

Drążkowska, J. and Y. Alibert (Dec. 2017). "Planetesimal formation starts at the snow line". In: *A&A* 608, A92, A92. DOI: 10.1051/0004-6361/201731491. arXiv: 1710.00009 [astro-ph.EP].

Dumusque, X. (Nov. 2018). "Measuring precise radial velocities on individual spectral lines. I. Validation of the method and application to mitigate stellar activity". In: *A&A* 620, A47, A47. DOI: 10.1051/0004-6361/201833795. arXiv: 1809.01548 [astro-ph.SR].

Dumusque, X. et al. (Jan. 2011). "Planetary detection limits taking into account stellar noise. I. Observational strategies to reduce stellar oscillation and granulation effects". In: *A&A* 525, A140, A140. DOI: 10.1051/0004-6361/201014097. arXiv: 1010.2616 [astro-ph.EP].

Eaton, J. A., G. W. Henry, and F. C. Fekel (June 2003). "Advantages of Automated Observing with Small Telescopes". In: *Astrophysics and Space Science Library*. Ed. by T. D. Oswalt. Vol. 288. Astrophysics and Space Science Library, p. 189. DOI: 10.1007/978-94-010-0253-0_38.

Endl, M., E. J. Brugamyer, et al. (Feb. 2016). “Two New Long-period Giant Planets from the McDonald Observatory Planet Search and Two Stars with Long-period Radial Velocity Signals Related to Stellar Activity Cycles”. In: *ApJ* 818.1, 34, p. 34. DOI: 10.3847/0004-637X/818/1/34. arXiv: 1512.02965 [astro-ph.EP].

Endl, M., W. D. Cochran, et al. (Sept. 2006). “Exploring the Frequency of Close-in Jovian Planets around M Dwarfs”. In: *ApJ* 649.1, pp. 436–443. DOI: 10.1086/506465. arXiv: astro-ph/0606121 [astro-ph].

Fabrycky, D. C. et al. (Aug. 2014). “Architecture of Kepler’s Multi-transiting Systems. II. New Investigations with Twice as Many Candidates”. In: *ApJ* 790, 146, p. 146. DOI: 10.1088/0004-637X/790/2/146. arXiv: 1202.6328 [astro-ph.EP].

Fang, J. and J.-L. Margot (Dec. 2012). “Architecture of Planetary Systems Based on Kepler Data: Number of Planets and Coplanarity”. In: *ApJ* 761, 92, p. 92. DOI: 10.1088/0004-637X/761/2/92. arXiv: 1207.5250 [astro-ph.EP].

Feng, Y. K. et al. (Feb. 2015). “The California Planet Survey IV: A Planet Orbiting the Giant Star HD 145934 and Updates to Seven Systems with Long-period Planets”. In: *ApJ* 800.1, 22, p. 22. DOI: 10.1088/0004-637X/800/1/22. arXiv: 1501.00633 [astro-ph.EP].

Fernandes, R. B. et al. (Mar. 2019). “Hints for a Turnover at the Snow Line in the Giant Planet Occurrence Rate”. In: *ApJ* 874.1, 81, p. 81. DOI: 10.3847/1538-4357/ab0300. arXiv: 1812.05569 [astro-ph.SR].

Fischer, D. A. and J. Valenti (Apr. 2005). “The Planet-Metallicity Correlation”. In: *ApJ* 622, pp. 1102–1117.

Fischer, D. A., G. Laughlin, et al. (Feb. 2005). “The N2K Consortium. I. A Hot Saturn Planet Orbiting HD 88133”. In: *ApJ* 620.1, pp. 481–486. DOI: 10.1086/426810.

Fischer, D. A., G. W. Marcy, and J. F. P. Spronck (Jan. 2014). “The Twenty-five Year Lick Planet Search”. In: *ApJS* 210.1, 5, p. 5. DOI: 10.1088/0067-0049/210/1/5. arXiv: 1310.7315 [astro-ph.EP].

Foreman-Mackey, D., D. W. Hogg, D. Lang, et al. (Mar. 2013a). “emcee: The MCMC Hammer”. In: *PASP* 125, p. 306. DOI: 10.1086/670067. arXiv: 1202.3665 [astro-ph.IM].

Foreman-Mackey, D., D. W. Hogg, D. Lang, et al. (Mar. 2013b). “emcee: The MCMC Hammer”. In: *PASP* 125.925, p. 306. DOI: 10.1086/670067. arXiv: 1202.3665 [astro-ph.IM].

Foreman-Mackey, D., D. W. Hogg, and T. D. Morton (Nov. 2014). “Exoplanet Population Inference and the Abundance of Earth Analogs from Noisy, Incomplete Catalogs”. In: *ApJ* 795.1, 64, p. 64. DOI: 10.1088/0004-637X/795/1/64. arXiv: 1406.3020 [astro-ph.EP].

Foreman-Mackey, D., T. D. Morton, et al. (Dec. 2016). “The Population of Long-period Transiting Exoplanets”. In: *AJ* 152.6, 206, p. 206. DOI: 10.3847/0004-6256/152/6/206. arXiv: 1607.08237 [astro-ph.EP].

Fulton, B. J., E. A. Petigura, A. W. Howard, et al. (Sept. 2017). “The California-Kepler Survey. III. A Gap in the Radius Distribution of Small Planets”. In: *AJ* 154, 109, p. 109. DOI: 10.3847/1538-3881/aa80eb. arXiv: 1703.10375 [astro-ph.EP].

Fulton, B. J., A. W. Howard, et al. (Oct. 2016). “Three Temperate Neptunes Orbiting Nearby Stars”. In: *ApJ* 830.1, 46, p. 46. DOI: 10.3847/0004-637X/830/1/46. arXiv: 1607.00007 [astro-ph.EP].

Fulton, B. J., E. A. Petigura, S. Blunt, et al. (Apr. 2018). “RadVel: The Radial Velocity Modeling Toolkit”. In: *PASP* 130.986, p. 044504. DOI: 10.1088/1538-3873/aaaaa8. arXiv: 1801.01947 [astro-ph.IM].

Fulton, B. J., L. J. Rosenthal, et al. (July 2021). “California Legacy Survey. II. Occurrence of Giant Planets beyond the Ice Line”. In: *ApJS* 255.1, 14, p. 14. DOI: 10.3847/1538-4365/abfcc1. arXiv: 2105.11584 [astro-ph.EP].

Gaia Collaboration, A. G. A. Brown, A. Vallenari, T. Prusti, J. H. J. de Bruijne, C. Babusiaux, et al. (Aug. 2018). “Gaia Data Release 2. Summary of the contents and survey properties”. In: *A&A* 616, A1, A1. DOI: 10.1051/0004-6361/201833051. arXiv: 1804.09365 [astro-ph.GA].

Gaia Collaboration, A. G. A. Brown, A. Vallenari, T. Prusti, J. H. J. de Bruijne, F. Mignard, et al. (Nov. 2016). “Gaia Data Release 1. Summary of the astrometric, photometric, and survey properties”. In: *A&A* 595, A2, A2. DOI: 10.1051/0004-6361/201629512. arXiv: 1609.04172 [astro-ph.IM].

Gibson, S. R. et al. (July 2018). “Keck Planet Finder: preliminary design”. In: *Ground-based and Airborne Instrumentation for Astronomy VII*. Ed. by C. J. Evans, L. Simard, and H. Takami. Vol. 10702. Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, p. 107025X. DOI: 10.1117/12.2311565.

Gonzalez, G. (Feb. 1997). “The stellar metallicity-giant planet connection”. In: *MNRAS* 285.2, pp. 403–412. DOI: 10.1093/mnras/285.2.403.

Hall, R. D. et al. (Sept. 2018). “On the Feasibility of Intense Radial Velocity Surveys for Earth-Twin Discoveries”. In: *MNRAS* 479.3, pp. 2968–2987. DOI: 10.1093/mnras/sty1464. arXiv: 1806.00518 [astro-ph.EP].

Hallatt, T. and E. J. Lee (Dec. 2020). “Can Large-scale Migration Explain the Giant Planet Occurrence Rate?” In: *ApJ* 904.2, 134, p. 134. DOI: 10.3847/1538-4357/abc1d7. arXiv: 2010.07949 [astro-ph.EP].

Hasegawa, Y. and R. E. Pudritz (Oct. 2011). “The origin of planetary system architectures - I. Multiple planet traps in gaseous discs”. In: *MNRAS* 417.2, pp. 1236–1259. DOI: 10.1111/j.1365-2966.2011.19338.x. arXiv: 1105.4015 [astro-ph.EP].

Haywood, R. D. et al. (Sept. 2014). “Planets and stellar activity: hide and seek in the CoRoT-7 system”. In: *MNRAS* 443.3, pp. 2517–2531. DOI: 10.1093/mnras/stu1320. arXiv: 1407.1044 [astro-ph.EP].

He, M. Y., E. B. Ford, and D. Ragozzine (Dec. 2019). “Architectures of exoplanetary systems - I. A clustered forward model for exoplanetary systems around Kepler’s FGK stars”. In: *MNRAS* 490.4, pp. 4575–4605. DOI: 10.1093/mnras/stz2869. arXiv: 1907.07773 [astro-ph.EP].

He, M. Y., E. B. Ford, D. Ragozzine, and D. Carrera (Dec. 2020). “Architectures of Exoplanetary Systems. III. Eccentricity and Mutual Inclination Distributions of AMD-stable Planetary Systems”. In: *AJ* 160.6, 276, p. 276. DOI: 10.3847/1538-3881/abba18. arXiv: 2007.14473 [astro-ph.EP].

Henry, G. W. (July 1999). “Techniques for Automated High-Precision Photometry of Sun-like Stars”. In: *PASP* 111, pp. 845–860. DOI: 10.1086/316388.

Henry, G. W., S. R. Kane, et al. (May 2013). “Host Star Properties and Transit Exclusion for the HD 38529 Planetary System”. In: *ApJ* 768.2, 155, p. 155. DOI: 10.1088/0004-637X/768/2/155. arXiv: 1303.4735 [astro-ph.EP].

Henry, G. W., G. W. Marcy, et al. (Jan. 2000). “A Transiting “51 Peg-like” Planet”. In: *ApJL* 529.1, pp. L41–L44. DOI: 10.1086/312458.

Hillenbrand, L. et al. (Jan. 2015). “Empirical Limits on Radial Velocity - Planet Detection for Young Stars”. In: *18th Cambridge Workshop on Cool Stars, Stellar Systems, and the Sun*. Vol. 18. Cambridge Workshop on Cool Stars, Stellar Systems, and the Sun, pp. 759–766. arXiv: 1408.3475 [astro-ph.SR].

Hirsch, L. A. et al. (Mar. 2021). “Understanding the Impacts of Stellar Companions on Planet Formation and Evolution: A Survey of Stellar and Planetary Companions within 25 pc”. In: *AJ* 161.3, 134, p. 134. DOI: 10.3847/1538-3881/abd639. arXiv: 2012.09190 [astro-ph.EP].

Hogg, D. W., A. D. Myers, and J. Bovy (Dec. 2010). “Inferring the Eccentricity Distribution”. In: *ApJ* 725.2, pp. 2166–2175. DOI: 10.1088/0004-637X/725/2/2166. arXiv: 1008.4146 [astro-ph.SR].

Holczer, T. et al. (July 2016). “Transit Timing Observations from Kepler. IX. Catalog of the Full Long-cadence Data Set”. In: *ApJS* 225.1, 9, p. 9. DOI: 10.3847/0067-0049/225/1/9. arXiv: 1606.01744 [astro-ph.EP].

Hopkins, P. F. and J. L. Christiansen (Oct. 2013). “Turbulent Disks are Never Stable: Fragmentation and Turbulence-promoted Planet Formation”. In: *ApJ* 776.1, 48, p. 48. DOI: 10.1088/0004-637X/776/1/48. arXiv: 1301.2600 [astro-ph.EP].

Horne, J. H. and S. L. Baliunas (Mar. 1986). "A Prescription for Period Analysis of Unevenly Sampled Time Series". In: *ApJ* 302, p. 757. DOI: 10.1086/164037.

Hoskin, M. (2003). *The History of Astronomy: A Very Short Introduction*. DOI: 10.1093/actrade/9780192803061.001.0001.

Howard, A. W., G. W. Marcy, S. T. Bryson, et al. (Aug. 2012). "Planet Occurrence within 0.25 AU of Solar-type Stars from Kepler". In: *ApJS* 201, 15, p. 15. DOI: 10.1088/0067-0049/201/2/15. arXiv: 1103.2541 [astro-ph.EP].

Howard, A. W. and B. J. Fulton (Nov. 2016). "Limits on Planetary Companions from Doppler Surveys of Nearby Stars". In: *PASP* 128.969, p. 114401. DOI: 10.1088/1538-3873/128/969/114401. arXiv: 1606.03134 [astro-ph.EP].

Howard, A. W., J. A. Johnson, G. W. Marcy, D. A. Fischer, J. T. Wright, D. Bernat, et al. (Oct. 2010). "The California Planet Survey. I. Four New Giant Exoplanets". In: *ApJ* 721.2, pp. 1467–1481. DOI: 10.1088/0004-637X/721/2/1467. arXiv: 1003.3488 [astro-ph.EP].

Howard, A. W., J. A. Johnson, G. W. Marcy, D. A. Fischer, J. T. Wright, G. W. Henry, et al. (Jan. 2011). "The NASA-UC Eta-Earth Program. II. A Planet Orbiting HD 156668 with a Minimum Mass of Four Earth Masses". In: *ApJ* 726.2, 73, p. 73. DOI: 10.1088/0004-637X/726/2/73. arXiv: 1003.3444 [astro-ph.EP].

Howard, A. W., G. W. Marcy, J. A. Johnson, et al. (Oct. 2010). "The Occurrence and Mass Distribution of Close-in Super-Earths, Neptunes, and Jupiters". In: *Science* 330.6004, p. 653. DOI: 10.1126/science.1194854. arXiv: 1011.0143 [astro-ph.EP].

Huang, C., Y. Wu, and A. H. M. J. Triaud (July 2016). "Warm Jupiters Are Less Lonely than Hot Jupiters: Close Neighbors". In: *ApJ* 825.2, 98, p. 98. DOI: 10.3847/0004-637X/825/2/98. arXiv: 1601.05095 [astro-ph.EP].

Huber, D. (May 2017). *Isoclassify: V1.2*. Version v1.2. DOI: 10.5281/zenodo.573372.

Hunter, J. D. (2007). "Matplotlib: A 2D graphics environment". In: *Computing In Science & Engineering* 9.3, pp. 90–95. DOI: 10.1109/MCSE.2007.55.

Ida, S. and D. N. C. Lin (Mar. 2004). "Toward a Deterministic Model of Planetary Formation. I. A Desert in the Mass and Semimajor Axis Distributions of Extrasolar Planets". In: *ApJ* 604, pp. 388–413.

Ida, S. and D. N. C. Lin (Jan. 2008a). "Toward a Deterministic Model of Planetary Formation. IV. Effects of Type I Migration". In: *ApJ* 673, pp. 487–501. DOI: 10.1086/523754. eprint: arXiv:0709.1375.

Ida, S. and D. N. C. Lin (Sept. 2008b). "Toward a Deterministic Model of Planetary Formation. V. Accumulation Near the Ice Line and Super-Earths". In: *ApJ* 685, pp. 584–595. DOI: 10.1086/590401.

Ida, S., H. Tanaka, et al. (Sept. 2018). “Slowing Down Type II Migration of Gas Giants to Match Observational Data”. In: *ApJ* 864.1, 77, p. 77. DOI: 10.3847/1538-4357/aad69c. arXiv: 1807.10871 [astro-ph.EP].

Isaacson, H. and D. Fischer (Dec. 2010). “Chromospheric Activity and Jitter Measurements for 2630 Stars on the California Planet Search”. In: *ApJ* 725.1, pp. 875–885. DOI: 10.1088/0004-637X/725/1/875. arXiv: 1009.2301 [astro-ph.EP].

Jackson, J. M. et al. (Apr. 2021). “Observable Predictions from Perturber-coupled High-eccentricity Tidal Migration of Warm Jupiters”. In: *AJ* 161.4, 200, p. 200. DOI: 10.3847/1538-3881/abe61f. arXiv: 2102.03427 [astro-ph.EP].

Jacob, W. S. (June 1855). “On certain Anomalies presented by the Binary Star 70 Ophiuchi”. In: *MNRAS* 15, p. 228. DOI: 10.1093/mnras/15.9.228.

Jennings, J., B. Ercolano, and G. P. Rosotti (July 2018). “The comparative effect of FUV, EUV and X-ray disc photoevaporation on gas giant separations”. In: *MNRAS* 477.3, pp. 4131–4141. DOI: 10.1093/mnras/sty964. arXiv: 1803.00571 [astro-ph.EP].

Johnson, J. A., K. M. Aller, et al. (Aug. 2010). “Giant Planet Occurrence in the Stellar Mass-Metallicity Plane”. In: *PASP* 122.894, p. 905. DOI: 10.1086/655775. arXiv: 1005.3084 [astro-ph.EP].

Johnson, J. A., B. P. Bowler, et al. (Oct. 2010). “A Hot Jupiter Orbiting the 1.7 M_{sun} Subgiant HD 102956”. In: *ApJL* 721.2, pp. L153–L157. DOI: 10.1088/2041-8205/721/2/L153. arXiv: 1007.4555 [astro-ph.EP].

Johnson, J. A., E. A. Petigura, et al. (Sept. 2017). “The California-Kepler Survey. II. Precise Physical Properties of 2025 Kepler Planets and Their Host Stars”. In: *AJ* 154.3, 108, p. 108. DOI: 10.3847/1538-3881/aa80e7. arXiv: 1703.10402 [astro-ph.EP].

Jones, M. I. et al. (May 2016). “Four new planets around giant stars and the mass-metallicity correlation of planet-hosting stars”. In: *A&A* 590, A38, A38. DOI: 10.1051/0004-6361/201628067. arXiv: 1603.03738 [astro-ph.EP].

Kane, S. R. et al. (June 2019). “Detection of Planetary and Stellar Companions to Neighboring Stars via a Combination of Radial Velocity and Direct Imaging Techniques”. In: *AJ* 157.6, 252, p. 252. DOI: 10.3847/1538-3881/ab1ddf. arXiv: 1904.12931 [astro-ph.EP].

Kant, I. (1755). *Allgemeine Naturgeschichte und Theorie des Himmels*. Konigsberg and Leipzig.

Kipping, D. M. (July 2013). “Parametrizing the exoplanet eccentricity distribution with the beta distribution.” In: *MNRAS* 434, pp. L51–L55. DOI: 10.1093/mnrasl/slt075. arXiv: 1306.4982 [astro-ph.EP].

Kipping, D. M. et al. (Apr. 2016). “A Transiting Jupiter Analog”. In: *ApJ* 820.2, 112, p. 112. DOI: 10.3847/0004-637X/820/2/112. arXiv: 1603.00042 [astro-ph.EP].

Kley, W. and R. P. Nelson (Sept. 2012). “Planet-Disk Interaction and Orbital Evolution”. In: *ARA&A* 50, pp. 211–249. DOI: 10.1146/annurev-astro-081811-125523. arXiv: 1203.1184 [astro-ph.EP].

Knutson, H. A. et al. (Apr. 2014). “Friends of Hot Jupiters. I. A Radial Velocity Search for Massive, Long-period Companions to Close-in Gas Giant Planets”. In: *ApJ* 785.2, 126, p. 126. DOI: 10.1088/0004-637X/785/2/126. arXiv: 1312.2954 [astro-ph.EP].

Kutra, T. and Y. Wu (Mar. 2020). “Kepler Planets and Metallicity”. In: *arXiv e-prints*, arXiv:2003.08431, arXiv:2003.08431. arXiv: 2003.08431 [astro-ph.EP].

Lee, E. J. (June 2019). “The Boundary between Gas-rich and Gas-poor Planets”. In: *ApJ* 878.1, 36, p. 36. DOI: 10.3847/1538-4357/ab1b40. arXiv: 1904.10470 [astro-ph.EP].

Lin, D. N. C. and J. Papaloizou (Oct. 1986). “On the Tidal Interaction between Protoplanets and the Protoplanetary Disk. III. Orbital Migration of Protoplanets”. In: *ApJ* 309, p. 846. DOI: 10.1086/164653.

Lindegren, L. et al. (Aug. 2018). “Gaia Data Release 2. The astrometric solution”. In: *A&A* 616, A2, A2. DOI: 10.1051/0004-6361/201832727. arXiv: 1804.09366 [astro-ph.IM].

Lissauer, J. J. et al. (Feb. 2009). “Models of Jupiter’s growth incorporating thermal and hydrodynamic constraints”. In: *Icarus* 199.2, pp. 338–350. DOI: 10.1016/j.icarus.2008.10.004. arXiv: 0810.5186 [astro-ph].

Lockwood, G. W. et al. (2013). “Decadal Variations of Sun-Like Stars”. In: *New Quests in Stellar Astrophysics III: A Panchromatic View of Solar-Like Stars, With and Without Planets*. Ed. by M. Chavez et al. Vol. 472. Astronomical Society of the Pacific Conference Series, p. 203.

Lu, P. K. et al. (Nov. 1987). “ICCD Speckle Observations of Binary Stars. III. A Survey for Duplicity Among High Velocity Stars”. In: *AJ* 94, p. 1318. DOI: 10.1086/114569.

Lucy, L. B. and M. A. Sweeney (Aug. 1971). “Spectroscopic binaries with circular orbits.” In: *AJ* 76, pp. 544–556. DOI: 10.1086/111159.

Ma, B. et al. (Oct. 2018). “The first super-Earth detection from the high cadence and high radial velocity precision Dharma Planet Survey”. In: *MNRAS* 480.2, pp. 2411–2422. DOI: 10.1093/mnras/sty1933. arXiv: 1807.07098 [astro-ph.EP].

Marcy, G. W., R. P. Butler, D. A. Fischer, et al. (Dec. 2002). “A Planet at 5 AU around 55 Cancri”. In: *ApJ* 581.2, pp. 1375–1388. DOI: 10.1086/344298. arXiv: astro-ph/0207294 [astro-ph].

Marcy, G. W., R. P. Butler, S. S. Vogt, et al. (Oct. 1998). “A Planetary Companion to a Nearby M4 Dwarf, Gliese 876”. In: *ApJL* 505.2, pp. L147–L149. DOI: 10.1086/311623. arXiv: astro-ph/9807307 [astro-ph].

Mayor, M., M. Marmier, et al. (Sept. 2011). “The HARPS search for southern extra-solar planets XXXIV. Occurrence, mass distribution and orbital properties of super-Earths and Neptune-mass planets”. In: *arXiv e-prints*, arXiv:1109.2497, arXiv:1109.2497. arXiv: 1109.2497 [astro-ph.EP].

Mayor, M., F. Pepe, et al. (Dec. 2003). “Setting New Standards with HARPS”. In: *The Messenger* 114, pp. 20–24.

Mayor, M. and D. Queloz (Nov. 1995). “A Jupiter-mass companion to a solar-type star”. In: *Nature* 378.6555, pp. 355–359. DOI: 10.1038/378355a0.

McKinney, W. (2010). “Data Structures for Statistical Computing in Python”. In: *Proceedings of the 9th Python in Science Conference*. Ed. by S. van der Walt and J. Millman, pp. 51–56.

Millholland, S., S. Wang, and G. Laughlin (Nov. 2017). “Kepler Multi-planet Systems Exhibit Unexpected Intra-system Uniformity in Mass and Radius”. In: *ApJL* 849.2, L33, p. L33. DOI: 10.3847/2041-8213/aa9714. arXiv: 1710.11152 [astro-ph.EP].

Millholland, S. C. and J. N. Winn (Oct. 2021). “Split Peas in a Pod: Intra-system Uniformity of Super-Earths and Sub-Neptunes”. In: *ApJL* 920.2, L34, p. L34. DOI: 10.3847/2041-8213/ac2c77. arXiv: 2110.01466 [astro-ph.EP].

Moe, M. and K. M. Kratter (Dec. 2019). “Impact of Binary Stars on Planet Statistics – I. Planet Occurrence Rates, Trends with Stellar Mass, and Wide Companions to Hot Jupiter Hosts”. In: *arXiv e-prints*, arXiv:1912.01699, arXiv:1912.01699. arXiv: 1912.01699 [astro-ph.EP].

Montet, B. T. et al. (Jan. 2014). “The TRENDS High-contrast Imaging Survey. IV. The Occurrence Rate of Giant Planets around M Dwarfs”. In: *ApJ* 781.1, 28, p. 28. DOI: 10.1088/0004-637X/781/1/28. arXiv: 1307.5849 [astro-ph.EP].

Morbidelli, A. et al. (Sept. 2015). “The great dichotomy of the Solar System: Small terrestrial embryos and massive giant planet cores”. In: *Icarus* 258, pp. 418–429. DOI: 10.1016/j.icarus.2015.06.003. arXiv: 1506.01666 [astro-ph.EP].

Mordasini, C., Y. Alibert, and W. Benz (July 2009). “Extrasolar planet population synthesis. I. Method, formation tracks, and mass-distance distribution”. In: *A&A* 501, pp. 1139–1160. DOI: 10.1051/0004-6361/200810301. arXiv: 0904.2524 [astro-ph.EP].

Mordasini, C. (2018). “Planetary Population Synthesis”. In: *Handbook of Exoplanets*. Springer International Publishing AG, p. 143. DOI: 10.1007/978-3-319-55333-7_143.

Moriarty, J. and D. Fischer (Aug. 2015). “Building Massive Compact Planetesimal Disks from the Accretion of Pebbles”. In: *ApJ* 809.1, 94, p. 94. DOI: 10.1088/0004-637X/809/1/94. arXiv: 1507.08215 [astro-ph.EP].

Morton, T. D. et al. (May 2016). "False Positive Probabilities for all Kepler Objects of Interest: 1284 Newly Validated Planets and 428 Likely False Positives". In: *ApJ* 822, 86, p. 86. DOI: 10.3847/0004-637X/822/2/86. arXiv: 1605.02825 [astro-ph.EP].

Naef, D. et al. (Mar. 2005). "The ELODIE planet search: synthetic view of the survey and its global detection threshold". In: *13th Cambridge Workshop on Cool Stars, Stellar Systems and the Sun*. Ed. by F. Favata, G. A. J. Hussain, and B. Battrick. Vol. 560. ESA Special Publication, p. 833. arXiv: astro-ph/0409230 [astro-ph].

Naoz, S. (Sept. 2016). "The Eccentric Kozai-Lidov Effect and Its Applications". In: *ARA&A* 54, pp. 441–489. DOI: 10.1146/annurev-astro-081915-023315. arXiv: 1601.07175 [astro-ph.EP].

Neil, A. R. and L. A. Rogers (Mar. 2020). "A Joint Mass-Radius-Period Distribution of Exoplanets". In: *ApJ* 891.1, 12, p. 12. DOI: 10.3847/1538-4357/ab6a92. arXiv: 1911.03582 [astro-ph.EP].

Nelson, B. E. et al. (Jan. 2016). "An empirically derived three-dimensional Laplace resonance in the Gliese 876 planetary system". In: *MNRAS* 455.3, pp. 2484–2499. DOI: 10.1093/mnras/stv2367. arXiv: 1504.07995 [astro-ph.EP].

Nielsen, E. L. et al. (July 2019). "The Gemini Planet Imager Exoplanet Survey: Giant Planet and Brown Dwarf Demographics from 10 to 100 au". In: *AJ* 158.1, 13, p. 13. DOI: 10.3847/1538-3881/ab16e9. arXiv: 1904.05358 [astro-ph.EP].

Ogilvie, G. I. (Aug. 2014). "Tidal Dissipation in Stars and Giant Planets". In: *ARA&A* 52, pp. 171–210. DOI: 10.1146/annurev-astro-081913-035941. arXiv: 1406.2207 [astro-ph.SR].

Ormel, C. W., B. Liu, and D. Schoonenberg (July 2017). "Formation of TRAPPIST-1 and other compact systems". In: *A&A* 604, A1, A1. DOI: 10.1051/0004-6361/201730826. arXiv: 1703.06924 [astro-ph.EP].

Pepper, J. (Jan. 2007). "KELT: The Kilodegree Extremely Little Telescope". PhD thesis. The Ohio State University.

Petigura, E. A. (Oct. 2015). "Prevalence of Earth-size Planets Orbiting Sun-like Stars". In: *arXiv:1510.03902*. arXiv: 1510.03902 [astro-ph.EP].

Petigura, E. A., A. W. Howard, and G. W. Marcy (Nov. 2013). "Prevalence of Earth-size planets orbiting Sun-like stars". In: *Proceedings of the National Academy of Science* 110, pp. 19273–19278. arXiv: 1311.6806 [astro-ph.EP].

Petigura, E. A., G. W. Marcy, and A. W. Howard (June 2013). "A Plateau in the Planet Population below Twice the Size of Earth". In: *ApJ* 770, 69, p. 69. DOI: 10.1088/0004-637X/770/1/69. arXiv: 1304.0460 [astro-ph.EP].

Petigura, E. A., G. W. Marcy, J. N. Winn, et al. (Feb. 2018). "The California-Kepler Survey. IV. Metal-rich Stars Host a Greater Diversity of Planets". In: *AJ* 155.2, 89, p. 89. DOI: 10.3847/1538-3881/aaa54c. arXiv: 1712.04042 [astro-ph.EP].

Petrovich, C. and S. Tremaine (Oct. 2016). "Warm Jupiters from Secular Planet-Planet Interactions". In: *ApJ* 829.2, 132, p. 132. DOI: 10.3847/0004-637X/829/2/132. arXiv: 1604.00010 [astro-ph.EP].

Pollacco, D. L. et al. (Oct. 2006). "The WASP Project and the SuperWASP Cameras". In: *PASP* 118.848, pp. 1407–1418. DOI: 10.1086/508556. arXiv: astro-ph/0608454 [astro-ph].

Pollack, J. B. et al. (Nov. 1996). "Formation of the Giant Planets by Concurrent Accretion of Solids and Gas". In: *Icarus* 124.1, pp. 62–85. DOI: 10.1006/icar.1996.0190.

Price-Whelan, A. M. et al. (Mar. 2017). "The Joker: A Custom Monte Carlo Sampler for Binary-star and Exoplanet Radial Velocity Data". In: *ApJ* 837.1, 20, p. 20. DOI: 10.3847/1538-4357/aa5e50. arXiv: 1610.07602 [astro-ph.SR].

Radovan, M. V. et al. (2014). "The automated planet finder at Lick Observatory". In: *Proceedings of the SPIE, Volume 9145, id. 91452B 12 pp. (2014).* Vol. 9145. Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, 91452B. DOI: 10.1117/12.2057310.

Rajpaul, V. et al. (Sept. 2015). "A Gaussian process framework for modelling stellar activity signals in radial velocity data". In: *MNRAS* 452.3, pp. 2269–2291. DOI: 10.1093/mnras/stv1428. arXiv: 1506.07304 [astro-ph.EP].

Raymond, S. N. and A. Morbidelli (2022). "Planet Formation: Key Mechanisms and Global Models". In: *Demographics of Exoplanetary Systems: Lecture Notes of the 3rd Advanced School on Exoplanetary Science*. Ed. by K. Biazzo et al. Cham: Springer International Publishing, pp. 3–82. ISBN: 978-3-030-88124-5. DOI: 10.1007/978-3-030-88124-5_1. URL: https://doi.org/10.1007/978-3-030-88124-5_1.

Reffert, S. et al. (Feb. 2015). "Precise radial velocities of giant stars. VII. Occurrence rate of giant extrasolar planets as a function of mass and metallicity". In: *A&A* 574, A116, A116. DOI: 10.1051/0004-6361/201322360. arXiv: 1412.4634 [astro-ph.EP].

Ricker, G. R. et al. (Aug. 2014). "Transiting Exoplanet Survey Satellite (TESS)". In: *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*. Vol. 9143. Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, p. 20. DOI: 10.1117/12.2063489. arXiv: 1406.0151 [astro-ph.EP].

Ricker, G. R. et al. (Jan. 2015). "Transiting Exoplanet Survey Satellite (TESS)". In: *Journal of Astronomical Telescopes, Instruments, and Systems* 1, 014003, p. 014003. DOI: 10.1117/1.JATIS.1.1.014003.

Rickman, E. L. et al. (May 2019). "The CORALIE survey for southern extrasolar planets. XVIII. Three new massive planets and two low-mass brown dwarfs at greater than 5 AU separation". In: *A&A* 625, A71, A71. DOI: 10.1051/0004-6361/201935356. arXiv: 1904.01573 [astro-ph.EP].

Roberts Lewis C., J. et al. (Nov. 2011). "Know the Star, Know the Planet. I. Adaptive Optics of Exoplanet Host Stars". In: *AJ* 142.5, 175, p. 175. DOI: 10.1088/0004-6256/142/5/175. arXiv: 1109.4320 [astro-ph.SR].

Robertson, P., M. Endl, et al. (Apr. 2012). "The McDonald Observatory Planet Search: New Long-period Giant Planets and Two Interacting Jupiters in the HD 155358 System". In: *ApJ* 749.1, 39, p. 39. DOI: 10.1088/0004-637X/749/1/39. arXiv: 1202.0265 [astro-ph.EP].

Robertson, P., S. Mahadevan, et al. (July 2014). "Stellar activity masquerading as planets in the habitable zone of the M dwarf Gliese 581". In: *Science* 345.6195, pp. 440–444. DOI: 10.1126/science.1253253. arXiv: 1407.1049 [astro-ph.EP].

Rodigas, T. J. et al. (May 2011). "Direct Imaging Constraints on the Putative Exoplanet 14 Her C". In: *ApJ* 732.1, 10, p. 10. DOI: 10.1088/0004-637X/732/1/10. arXiv: 1102.3691 [astro-ph.EP].

Rosenthal, L. J., B. J. Fulton, et al. (July 2021). "The California Legacy Survey. I. A Catalog of 178 Planets from Precision Radial Velocity Monitoring of 719 Nearby Stars over Three Decades". In: *ApJS* 255.1, 8, p. 8. DOI: 10.3847/1538-4365/abe23c. arXiv: 2105.11583 [astro-ph.EP].

Rosenthal, L. J., H. A. Knutson, et al. (Dec. 2021). "The California Legacy Survey III. On The Shoulders of (Some) Giants: The Relationship between Inner Small Planets and Outer Massive Planets". In: *arXiv e-prints*, arXiv:2112.03399, arXiv:2112.03399. arXiv: 2112.03399 [astro-ph.EP].

Rowan, D. et al. (Feb. 2016). "The Lick-Carnegie Exoplanet Survey: HD 32963—A New Jupiter Analog Orbiting a Sun-like Star". In: *ApJ* 817.2, 104, p. 104. DOI: 10.3847/0004-637X/817/2/104. arXiv: 1512.00417 [astro-ph.EP].

Santos, N. C., G. Israelian, and M. Mayor (Mar. 2004). "Spectroscopic [Fe/H] for 98 extra-solar planet-host stars. Exploring the probability of planet formation". In: *A&A* 415, pp. 1153–1166.

Schlecker, M. et al. (July 2020). "The New Generation Planetary Population Synthesis (NGPPS). III. Warm super-Earths and cold Jupiters: A weak occurrence correlation, but with a strong architecture-composition link". In: *arXiv e-prints*, arXiv:2007.05563, arXiv:2007.05563. arXiv: 2007.05563 [astro-ph.EP].

Schoonenberg, D. and C. W. Ormel (June 2017). "Planetesimal formation near the snowline: in or out?" In: *A&A* 602, A21, A21. DOI: 10.1051/0004-6361/201630013. arXiv: 1702.02151 [astro-ph.EP].

Seager, S. and J. J. Lissauer (2010). "Introduction to Exoplanets". In: *Exoplanets*. Ed. by S. Seager, pp. 3–13.

Stassun, K. G., K. A. Collins, and B. S. Gaudi (Mar. 2017). “Accurate Empirical Radii and Masses of Planets and Their Host Stars with Gaia Parallaxes”. In: *AJ* 153.3, 136, p. 136. DOI: 10.3847/1538-3881/aa5df3. arXiv: 1609.04389 [astro-ph.EP].

Struve, O. (Oct. 1952). “Proposal for a project of high-precision stellar radial velocity work”. In: *The Observatory* 72, pp. 199–200.

Tal-Or, L. et al. (Mar. 2019). “Correcting HIRES/Keck radial velocities for small systematic errors”. In: *MNRAS* 484.1, pp. L8–L13. DOI: 10.1093/mnrasl/sly227. arXiv: 1810.02986 [astro-ph.EP].

Tamuz, O. et al. (Mar. 2008). “The CORALIE survey for southern extra-solar planets. XV. Discovery of two eccentric planets orbiting <ASTROBJ>HD 4113</ASTROBJ> and <ASTROBJ>HD 156846</ASTROBJ>”. In: *A&A* 480.3, pp. L33–L36. DOI: 10.1051/0004-6361:20078737. arXiv: 0710.5028 [astro-ph].

Tange, O. (Feb. 2011). “GNU Parallel - The Command-Line Power Tool”. In: *;login: The USENIX Magazine* 36.1, pp. 42–47. DOI: 10.5281/zenodo.16303. URL: http://www.gnu.org/s/parallel.

Thompson, S. J. et al. (Aug. 2016). “HARPS3 for a roboticized Isaac Newton Telescope”. In: *Ground-based and Airborne Instrumentation for Astronomy VI*. Ed. by C. J. Evans, L. Simard, and H. Takami. Vol. 9908. Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, 99086F. DOI: 10.1117/12.2232111. arXiv: 1608.04611 [astro-ph.IM].

Thorngren, D. P. et al. (Nov. 2016). “The Mass-Metallicity Relation for Giant Planets”. In: *ApJ* 831.1, 64, p. 64. DOI: 10.3847/0004-637X/831/1/64. arXiv: 1511.07854 [astro-ph.EP].

Tinney, C. G. et al. (Apr. 2001). “First Results from the Anglo-Australian Planet Search: A Brown Dwarf Candidate and a 51 Peg-like Planet”. In: *ApJ* 551.1, pp. 507–511. DOI: 10.1086/320097. arXiv: astro-ph/0012204 [astro-ph].

Trifonov, T. et al. (Apr. 2020). “Public HARPS radial velocity database corrected for systematic errors”. In: *A&A* 636, A74, A74. DOI: 10.1051/0004-6361/201936686. arXiv: 2001.05942 [astro-ph.EP].

Udry, S. et al. (May 2017). “The HARPS search for southern extra-solar planets. XXXVI. Eight HARPS multi-planet systems hosting 20 super-Earth and Neptune-mass companions”. In: *arXiv e-prints*, arXiv:1705.05153, arXiv:1705.05153. arXiv: 1705.05153 [astro-ph.EP].

Valenti, J. A., R. P. Butler, and G. W. Marcy (Oct. 1995). “Determining Spectrometer Instrumental Profiles Using FTS Reference Spectra”. In: *PASP* 107, p. 966. DOI: 10.1086/133645.

Van Eylen, V. et al. (Feb. 2019). “The Orbital Eccentricity of Small Planet Systems”. In: *AJ* 157.2, 61, p. 61. DOI: 10.3847/1538-3881/aaf22f. arXiv: 1807.00549 [astro-ph.EP].

Vogt, S. S., S. L. Allen, et al. (1994). "HIRES: the high-resolution echelle spectrometer on the Keck 10-m Telescope". In: *Proc. SPIE Instrumentation in Astronomy VIII, David L. Crawford; Eric R. Craine; Eds., Volume 2198, p. 362*. Ed. by D. L. Crawford and E. R. Craine. Vol. 2198. Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, p. 362. DOI: 10.1117/12.176725.

Vogt, S. S., M. Radovan, et al. (Apr. 2014). "APF-The Lick Observatory Automated Planet Finder". In: *PASP* 126, pp. 359–379. DOI: 10.1086/676120. arXiv: 1402.6684 [astro-ph.IM].

Vogt, S. S., R. P. Butler, J. Burt, et al. (Nov. 2017). "A Six-planet System around the Star HD 34445". In: *AJ* 154.5, 181, p. 181. DOI: 10.3847/1538-3881/aa8b61. arXiv: 1710.07337 [astro-ph.EP].

Vogt, S. S., R. P. Butler, G. W. Marcy, et al. (Mar. 2002). "Ten Low-Mass Companions from the Keck Precision Velocity Survey". In: *ApJ* 568.1, pp. 352–362. DOI: 10.1086/338768. arXiv: astro-ph/0110378 [astro-ph].

Vogt, S. S., R. A. Wittenmyer, et al. (Jan. 2010). "A Super-Earth and Two Neptunes Orbiting the Nearby Sun-like Star 61 Virginis". In: *ApJ* 708.2, pp. 1366–1375. DOI: 10.1088/0004-637X/708/2/1366. arXiv: 0912.2599 [astro-ph.EP].

Walt, S. van der, S. C. Colbert, and G. Varoquaux (Mar. 2011). "The NumPy Array: A Structure for Efficient Numerical Computation". In: *Computing in Science Engineering* 13.2, pp. 22–30. ISSN: 1521-9615. DOI: 10.1109/MCSE.2011.37.

Weiss, L. M. et al. (Jan. 2018). "The California-Kepler Survey. V. Peas in a Pod: Planets in a Kepler Multi-planet System Are Similar in Size and Regularly Spaced". In: *AJ* 155.1, 48, p. 48. DOI: 10.3847/1538-3881/aa9ff6. arXiv: 1706.06204 [astro-ph.EP].

Wittenmyer, R. A., R. P. Butler, et al. (Feb. 2020). "The Pan-Pacific Planet Search - VIII. Complete results and the occurrence rate of planets around low-luminosity giants". In: *MNRAS* 491.4, pp. 5248–5257. DOI: 10.1093/mnras/stz3378. arXiv: 1911.11954 [astro-ph.EP].

Wittenmyer, R. A., M. Endl, and W. D. Cochran (Jan. 2007). "Long-Period Objects in the Extrasolar Planetary Systems 47 Ursae Majoris and 14 Herculis". In: *ApJ* 654.1, pp. 625–632. DOI: 10.1086/509110. arXiv: astro-ph/0609117 [astro-ph].

Wittenmyer, R. A., M. Endl, W. D. Cochran, and H. F. Levison (Sept. 2007). "Dynamical and Observational Constraints on Additional Planets in Highly Eccentric Planetary Systems". In: *AJ* 134.3, pp. 1276–1284. DOI: 10.1086/520880. arXiv: 0706.1962 [astro-ph].

Wittenmyer, R. A., J. Horner, et al. (Apr. 2017). "The Anglo-Australian Planet Search. XXV. A Candidate Massive Saturn Analog Orbiting HD 30177". In: *AJ* 153.4, 167, p. 167. DOI: 10.3847/1538-3881/aa5f17. arXiv: 1612.02072 [astro-ph.EP].

Wittenmyer, R. A., S. Wang, et al. (Feb. 2020). “Cool Jupiters greatly outnumber their toasty siblings: occurrence rates from the Anglo-Australian Planet Search”. In: *MNRAS* 492.1, pp. 377–383. DOI: 10.1093/mnras/stz3436. arXiv: 1912.01821 [astro-ph.EP].

Wittrock, J. M. et al. (Nov. 2017). “Exclusion of Stellar Companions to Exoplanet Host Stars”. In: *AJ* 154.5, 184, p. 184. DOI: 10.3847/1538-3881/aa8d69. arXiv: 1709.05315 [astro-ph.EP].

Wolszczan, A. and D. A. Frail (Jan. 1992). “A planetary system around the millisecond pulsar PSR1257 + 12”. In: *Nature* 355.6356, pp. 145–147. DOI: 10.1038/355145a0.

Wright, J. T., G. W. Marcy, R. P. Butler, et al. (Aug. 2008). “The Jupiter Twin HD 154345b”. In: *ApJL* 683.1, p. L63. DOI: 10.1086/587461. arXiv: 0802.1731 [astro-ph].

Wright, J. T., G. W. Marcy, D. A. Fischer, et al. (Mar. 2007). “Four New Exoplanets and Hints of Additional Substellar Companions to Exoplanet Host Stars”. In: *ApJ* 657.1, pp. 533–545. DOI: 10.1086/510553. arXiv: astro-ph/0611658 [astro-ph].

Wright, J. T., S. Upadhyay, et al. (Mar. 2009). “Ten New and Updated Multiplanet Systems and a Survey of Exoplanetary Systems”. In: *ApJ* 693.2, pp. 1084–1099. DOI: 10.1088/0004-637X/693/2/1084. arXiv: 0812.1582 [astro-ph].

Wright, J. T. (Mar. 2016). “Twenty Years of Precise Radial Velocities at Keck and Lick Observatories”. In: *arXiv e-prints*, arXiv:1603.08384, arXiv:1603.08384. arXiv: 1603.08384 [astro-ph.SR].

Yee, S. W., E. A. Petigura, and K. von Braun (Feb. 2017). “Precision Stellar Characterization of FGKM Stars using an Empirical Spectral Library”. In: *ApJ* 836.1, 77, p. 77. DOI: 10.3847/1538-4357/836/1/77. arXiv: 1701.00922 [astro-ph.SR].

Youdin, A. N. and J. Goodman (Feb. 2005). “Streaming Instabilities in Protoplanetary Disks”. In: *ApJ* 620.1, pp. 459–469. DOI: 10.1086/426895. arXiv: astro-ph/0409263 [astro-ph].

Yu, Q. and S. Tremaine (Mar. 2001). “Resonant Capture by Inward-migrating Planets”. In: *AJ* 121.3, pp. 1736–1740. DOI: 10.1086/319401. arXiv: astro-ph/0009255 [astro-ph].

Zhao, L. L. et al. (Apr. 2022). “The EXPRES Stellar Signals Project II. State of the Field in Disentangling Photospheric Velocities”. In: *AJ* 163.4, 171, p. 171. DOI: 10.3847/1538-3881/ac5176. arXiv: 2201.10639 [astro-ph.EP].

Zhu, W. and Y. Wu (Sept. 2018). “The Super Earth-Cold Jupiter Relations”. In: *AJ* 156.3, 92, p. 92. DOI: 10.3847/1538-3881/aad22a. arXiv: 1805.02660 [astro-ph.EP].